ence Library

MAR 17 1945

Vol. 35, No. 1 February, 1945

Journal of Experimental Psychology

EDITED BY

SAMUEL W. FERNBERGER, ON LEAVE
FRANCIS W. IRWIN, ACTING EDITOR
UNIVERSITY OF PENNSYLVANIA

CONSULTING EDITORS

CHARLES W. BRAY, *Princeton University;* ELMER K. CULLER, *University of Rochester;* CLARENCE H. GRAHAM, *Brown University;* JOY P. GUILFORD, *University of Southern California;* FRANCIS W. IRWIN, *University of Pennsylvania;* DONALD G. MARQUIS, *Yale University;* ARTHUR W. MELTON, *University of Missouri;* CARROLL C. PRATT, *Rutgers University.*

CONTENTS

Utrocular, or 'Which Eye,' Discrimination: S. SMITH 1

The Acquisition of a Trace Conditioned Response as a Function of the Magnitude of the Stimulus Trace: B. REYNOLDS 15

Trial and Error Learning with Differential Cues: H. E. JONES 31

Experimental Studies in Affective Processes: II. On the Quantification and Evaluation of 'Measured' Changes in Skin Resistance: E. A. HAGGARD .. 46

A Study of Individual Differences in Motion Acuity at Scotopic Levels of Illumination: C. J. WARDEN, H. C. BROWN, AND S. ROSS 57

Generalization in Memory: J. GOODWIN, L. LONG, AND L. WELCH 71

An Alternative Explanation of the Effect of Temperature upon Retention in the Goldfish: F. N. JONES 76

PUBLISHED BI-MONTHLY,
BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STS., LANCASTER, PA.
AND NORTHWESTERN UNIVERSITY, EVANSTON, ILLINOIS

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879

PUBLICATIONS OF

THE AMERICAN PSYCHOLOGICAL ASSOCIATION

WILLARD L. VALENTINE, *Business Manager*

PSYCHOLOGICAL REVIEW

HERBERT S. LANGFELD, *Editor*
Princeton University

Contains original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.
Subscription: $5.50 (Foreign, $5.75). Single copies, $1.00.

PSYCHOLOGICAL BULLETIN

JOHN E. ANDERSON, *Editor*
University of Minnesota

Contains critical reviews of books and articles, psychological news and notes, university notices, and announcements. Appears monthly (10 issues), the annual volume comprising about 665 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.
Subscription: $7.00 (Foreign, $7.25). Single copies, 75c.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

S. W. FERNBERGER, *Editor, on Leave,* FRANCIS W. IRWIN, *Acting Editor*
University of Pennsylvania

Contains original contributions of an experimental character. Appears bi-monthly (beginning 1944), one volume per year, each volume of six numbers containing about 520 pages.
Subscription: $7.00 (Foreign, $7.25). Single copies, $1.25.

PSYCHOLOGICAL ABSTRACTS

WALTER S. HUNTER, *Editor*
Brown University

Appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.
Subscription: $7.00 (Foreign, $7.25). Single copies, 75c.

PSYCHOLOGICAL MONOGRAPHS

JOHN F. DASHIELL, *Editor*
University of North Carolina

Consist of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 350 pages.
Subscription: $6.00 per volume (Foreign, $6.30).

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

GORDON W. ALLPORT, *Editor*
Harvard University

Appears quarterly, January, April, July, October, the four numbers comprising a volume of 400 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.
Subscription: $5.00 (Foreign, $5.25). Single copies, $1.50.

JOURNAL OF APPLIED PSYCHOLOGY

DONALD G. PATERSON, *Editor*
University of Minnesota

Covers the applications of psychology in business, industry, education, etc. Appears bi-monthly, February, April, June, August, October, and December.
Subscription: $6.00. Single copies, $1.25.

Subscriptions, orders, and business communications should be sent to

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

NORTHWESTERN UNIVERSITY, EVANSTON, ILLINOIS

Journal of
Experimental Psychology

VOL. 35, NO. 1 FEBRUARY, 1945

UTROCULAR, OR 'WHICH EYE,' DISCRIMINATION [1]

BY STEVENSON SMITH

University of Washington

The present experimental problem is this: can stimulation of the right retina be distinguished from stimulation of the left retina in the absence of all other sensory cues?

We have many means of telling which of our eyes has been stimulated by a light. We may bring into play sense organs, other than those in the retina, whose local sign is not to be questioned. If a flash-light touches the skin near the eye, or if we watch the flash-light approaching, we may owe our discrimination to tactile or kinesthetic stimuli. If a distant light in a dark room is allowed to stimulate an *S*'s right or left eye at random as a right or a left screen is removed from before the eye, the *S* may respond with a preponderance of correct answers just because his two eyes are likely to be converged upon a nearer point, and the light thus tends to fall upon the nasal halves of the two retinas. This may have occurred in the experiment of Thelin and Altman (**17**). The responses 'right' and 'left' may then really refer to the position of the light in the monocular field of vision. The light is displaced to the right in the projected field of the right eye and to the left in the field of the left eye, if involuntary convergence is upon some point nearer than the light and on a line between the light and the nose. In such a set-up a preponderance of correct responses cannot be regarded as evidence for utraretinal local sign. By utraretinal discrimination we shall mean a differential response to the stimulation of one or the other retina in the absence of any dependent stimulus variable.

[1] How shall we name that kind of discrimination that serves to designate that the right eye or the left eye has been stimulated? We cannot call it ocular discrimination, because this includes any form of discrimination in which an eye may engage. Neither is it binocular discrimination, for this only means that both eyes are employed. *Uter*, meaning which one of two, seems to be too precious a linguistic instrument to be neglected in the face of a real need, so in place of the clumsy form, *which-eye* discrimination, the writer begs leave to use the term utrocular discrimination.

1. Control of the Factor of Retinal Position of Stimulus

Experiment 1.—Results: When retinal position of stimulation is made independent of which eye it is that is stimulated, utrocular discrimination remains highly reliable.

In order to eliminate the dependent variable of more frequent stimulation of the nasal or temporal halves of the retinas, an apparatus was constructed that gave stimulation whose source was well to one side or the other of any probable line of regard. Each eye looked into one of two similar dark boxes. Sixteen pin-hole apertures, four being placed in a vertical line on each side of the far wall of each box, were illuminated in random order (Fig. 1). The light source was a

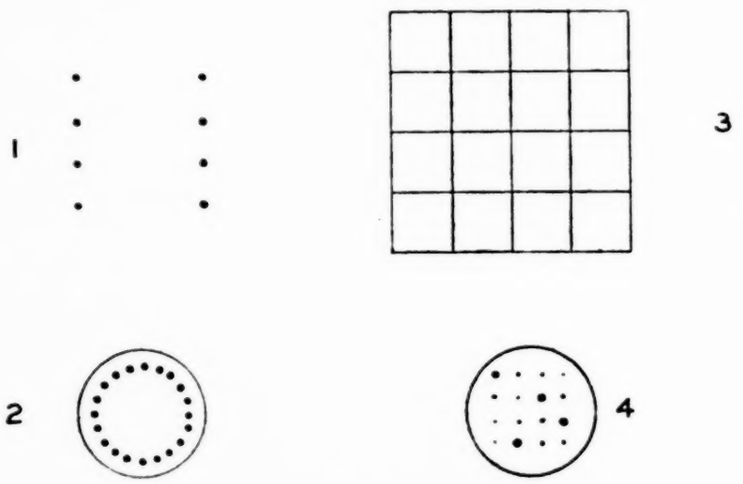

Fig. 1. These drawings represent the fixation patterns and stimulus patterns used in the eight experiments. In each case only one of the two duplicate patterns for the two eyes is here shown. *Pattern 1* was used in Experiments 1 and 2. Single point stimulation in randomized order. Background darkness. Eyes three cm. from plane of apertures. Horizontal distance between apertures two cm. (This establishes scale for all drawings.) *Pattern 2* used in Experiment 7. Single point stimulation or double point stimulation in randomized order. *Pattern 3* used in Experiments 3, 4, 5 and 6. Stimulus flash from center of either figure. Fixation at selected crosslines. *Pattern 4* used in Experiment 8. Only one pattern shown, seven others being obtained by 90 degree rotation and reflection. The four-point pattern is represented as seen binocularly. Two points in each pattern stimulate the right eye, the other two the left eye. Faint dots added to drawing to indicate nature of patterns.

neon bulb. Diffusion was secured by opal glass plates, the number of which also controlled brightness. Time of stimulation was .04 sec. The interval between stimulations was three sec.

Practice in responding to this situation resulted in marked improvement in discrimination. Armstrong (1) used 10 *S*s, each of whom gave 200 responses with this set-up, and secured 80 percent correct judgments. This proportion is 27 times the S.D. of chance fluctuation. There can be no doubt that utrocular discrimination occurs under these conditions. The writer, while serving as *S*, improved from a score of 58 correct judgments in 100 trials to 100 correct. It seemed thoroughly demonstrated in this preliminary experiment that, regardless of the position of the stimulus on the retina, the eye that was stimulated could be identified. But there are other dependent variables that had not been excluded.

In 1922 Guthrie and the present writer had suggested that local sign may be identical with the specific motor response which any localized stimulation innately calls out (**16**). Peterson (1926) in much the same way interpreted local signs as orientation tendencies

(13). If this view of local sign is correct, the cue to which we respond by saying 'right' or 'left' is not the light stimulus, but rather the stimulation resulting from eye muscle contractions which are themselves the more immediate results of the light. So considered, utrocular local sign becomes not an aspect of light sensation but a proprioceptive part of perception. We might go further and hold that there would be no local sign were there no original specific response tendency set off by the stimulation of a particular sense area. This original response tendency may become greatly elaborated and changed through conditioning, but it may well be essential to the ensuing elaboration.

The question was probably first clearly defined by Helmholtz in 1866 (10). He took the position that we cannot be certain which eye is stimulated and that our judgment will depend upon other sensory cues and upon learned interpretation. He never made it clear how two sense organs with only a single common response to begin with can individually through conditioning become linked selectively with separate responses. Eye movements were not for him one of the secondary considerations on which the judgment might depend. Schön (14) was almost alone in his belief in an indwelling quality that differentiates the visual sensations of the two eyes. The work of Fleisch (7) bore out Helmholtz's position and the results of Bourdon (2, 3), Heine (8, 9), and Brückner and Brücke (4, 5, 6) tended to confirm the main thesis. The last two experimenters, however, made much of the *scheinbares Organgefühl* that results from dark adapting one eye and then using both for ordinary vision. There is then not only a difference in sharpness and brightness between the images of the two eyes but also a feeling of discomfort which they believed belongs intrinsically to one eye. They proposed that such a feeling of faulty vision might result from a flash stimulation to a single eye and thus serve as the basis for discrimination. They repeated Heine's crude experiment, repeated again with modifications by Thelin and Altman (17), and were the first to report that objects briefly seen against an illuminated background were far more difficult to refer to the receiving eye than were sparks seen in the dark. It did not occur to them that an illuminated background affords fixation points which tend to anchor the eye and thus interfere with reflex movements upon which utrocular discrimination seems to depend.

2. Control of the Factor of Iris and Ciliary Muscle Accommodation

Introspectively there seemed to the writer to be a distinct kinesthetic stimulus cue to the discrimination in question, and this seemed to be located in the eye that was stimulated. The first attack was on

the sphincters of the eye. It is highly probable that the ipsilateral ciliary and iris reflexes are greater than the corresponding contralaterals. If so, the greater contraction on the side stimulated might furnish the essential cue for utrocular discrimination.

Experiment 2.—Results: The use of a cycloplegic did not reduce the degree of utrocular discrimination.

In order to find out whether the elimination of ciliary and iris proprioceptive cues would reduce the accuracy of discrimination, the writer used homatropin and cocaine in his own eyes immediately after he had given a series of 100 responses under normal conditions. In this series there were 98 right and 2 wrong responses. With what is clinically regarded as complete iris and ciliary relaxation there was no reduction whatever of his ability to discriminate. In the first series of 50 consecutive responses under conditions of iris and ciliary relaxation there were no errors.

The iris and ciliary muscles seem thus to be ruled out as necessary or even significant factors in utrocular discrimination. There remain, however, the external eye muscles. These, unfortunately, cannot be relaxed through the use of drugs as can the internal muscles. Some other method had to be devised to study their effect.

3. Control of the Factor of External Muscle Contraction

The blindfolded eye will go through much the same movements that the stimulated eye makes in following a light. Here the lateral movements of the two eyes are in the same direction but the internal rectus of one eye is coördinated with the external rectus of the other. This contralateral reflex is a muscle to muscle reflex and follows upon the ipsilateral reflex. The ipsilateral reflex movement of either eye is stronger and will triumph over its contralateral reflex if the two tendencies are incompatible. Thus, if the axes of vision are almost parallel and a near-by light stimulates the temporal halves of the two retinas, the ipsilateral reflex in each eye is convergence and the contralateral reflex in each eye is divergence. Convergence, i.e., the ipsilateral reflex, wins out. Thus we have a theoretical expectation of unequal proprioceptive stimulation in the two eyes when a small area of only one retina is stimulated and when fixation tendencies are consequently aroused. The implication is that utrocular discrimination may be a function of the inequality of action in the external muscles of the two eyes. If the reflexes of the stimulated and of the unstimulated eye differ in timing or in degree, the kinesthetic stimulus patterns from the two eyes must differ when only one eye is stimulated, and differing they must afford a proprioceptive cue for discriminating which retina it is that has been stimulated by light.

If, as seems to be true, each pair of corresponding points on the two retinas is always connected with a mutual area in the visual cortex, utraretinal local sign would seem to be impossible. This is to say, without the help of proprioception no utrocular discrimination could take place. There would be no difference in neural processes in the visual projection area, whether the right eye or the left eye were stimulated. We know that there are differences of eye muscle response to right and left eye visual stimulation, but such responses are mediated primarily by neural structures that are separately reserved for the right eye and the left eye. If the muscle sense organs in the right and the left eye have different projection areas, utrocular discrimination as well as stereoscopic depth perception may be primarily a function of proprioception. The light stimulation may determine only the point of departure, the extent, the direction, and the time of the eye movements.

Experiment 3.—Results: The greater the binocular convergence, as it varies from 5 to 25 degrees, the greater are the number of errors made in judging which eye has been stimulated. (Fig. 2.) In this experi-

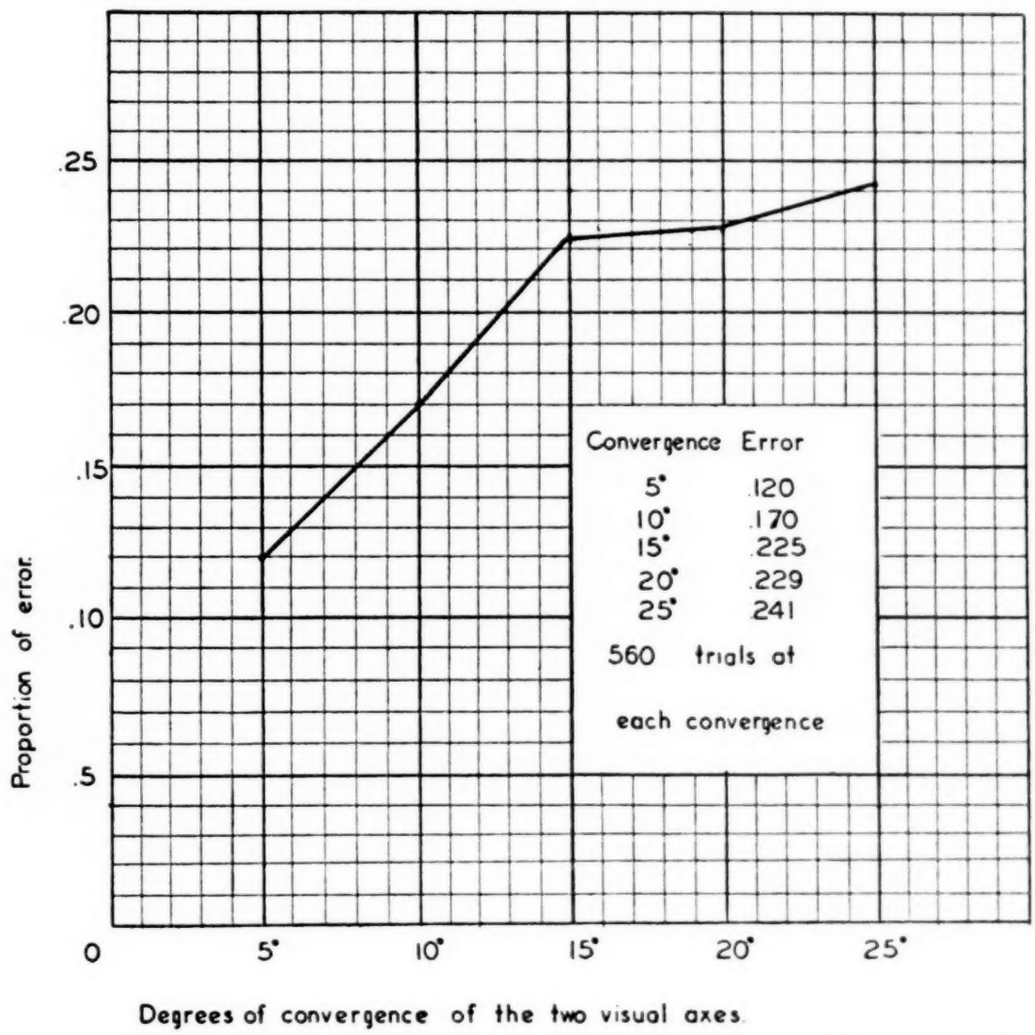

FIG. 2. Proportion of error as a function of degrees of convergence of the two visual axes.

ment foveal stimulation was used and the eyes were anchored on cross lines in an illuminated field (Fig. 1). The cross lines divided a square 1.52 in. on a side into 16 smaller squares. One such figure was 11 in. distant from each eye and screened from the other eye. The *S* obtained binocular fusion of these two similar figures with fixation at the center cross lines where a .02 in. aperture emitted a .05 ± .002 sec. stimulus flash from a diffused neon bulb to one eye or the other in randomized order every three sec. One and one-half sec. before each stimulation the figures were illuminated by a tungsten filament light that was extinguished with the beginning of stimulation. The rhythm enabled the *S* to obtain optimum fixation and binocular fusion at the moment the stimulus was given. The timing was controlled by a synchronous-motor-driven contact disc.

Equal intensity of stimulus to the two eyes was obtained empirically by giving the *S* a series of alternating right and left eye stimulations which, because of binocular fusion of the fields, appeared as a repeated flash from a single source. After the lights were balanced the stimulus boxes were interchanged and the lights, thus reversed, were again judged for apparent equality. This was done to discover any possible subjective difference between the right and left eye receptiveness that might have depended upon a structural or pathological difference in the two eyes. No subjective difference could be discovered. Nevertheless, at the end of each 40 trials during the experiment the stimulus boxes were half the time interchanged and half the time not interchanged, according to a randomized schedule. The *S* never knew whether the change had been made.

The nature of this experiment demands a trained *S*. If binocular fusion of the cross lines of the two figures is not precise, the resulting displacement of the stimulus upon the retina affords a cure to indicate which eye is stimulated. For the right eye an exophoria would displace the light to the left in the projected field. An esophoria would have the opposite effect. Dr. S. D. Spragg [2] acted as *E* and the writer served as *S*. All constant errors are believed to have been eliminated. Whenever a stimulus did not seem to the *S* to be directly foveal the trial was omitted and later added at the end of the 40-trial series.

Forty trials were given at each of the five degrees of convergence on 14 successive days. On odd days the degree of convergence was progressively increased from series to series and on even days it was decreased. Convergence was controlled by changing the distance between the stimulus boxes. Each lens assembly through which the light entered the eye contained the *S*'s refractive correction and a 15 diopter prism base out.

[2] Dr. Spragg also calibrated the timer.

The differences in accuracy of discrimination at various degrees of convergence are fairly valid throughout the range (Fig. 2). The error difference between 5 and 25 degrees of convergence is 5.4 times the S.D. of the difference. It seems clearly indicated that convergence interferes with judging which eye is stimulated. The muscle strain of convergence could interfere only with muscle cues, not with retinal cues. Hence, the results conform to the hypothesis that the rectus muscles provide specific cues for utrocular discrimination. Whether the retinas provide specific cues is doubtful. That question will be considered in the latter part of this article.

Experiment 4.—Results: Utrocular discrimination is better for foveal stimulation than for peripheral stimulation. Before Experiment 3 was run, the same *E* and *S* using the same set-up determined the difference in utrocular discrimination when foveal stimulation and when peripheral stimulation were used. There were 720 trials under each of these conditions. For six days 40 trials with foveal stimulation alternated with 40 trials with peripheral stimulation for a total of 240 trials a day. Foveal stimulation was secured as in the experiment just described. For peripheral stimulation the *S* fixated the cross lines that were .38 in. to the right of the light apertures. The distance from eye to fixation point was 11 in. and convergence was 15 degrees. With foveal stimulation there was a $.35 \pm .012$ error, and for peripheral stimulation there was a $.42 \pm .012$ error. The practice effect was comparable under the two conditions. The difference is 2.7 times the S.D. of the difference with a *P* of .997.

Under these conditions the judgment as to which eye is stimulated is more often correct when foveal stimulation is used than when peripheral stimulation is used. This fits the hypothesis that with monocular foveal stimulation there is a greater intensity difference between the rectus muscle cues received by the two eyes than there is with monocular peripheral stimulation. All the rectus muscles of the eye foveally stimulated are innervated. It is highly doubtful that there is a corresponding innervation of the rectus muscles of the other eye that is nearly as great. If there were it would interfere with the convergence accommodation of the other eye when a point of light stimulates one eye foveally and the other eye peripherally. But when one eye only is stimulated peripherally the other eye responds with a contralateral reflex which is in the same direction as the ipsilateral reflex. If there are fewer muscles innervated in peripheral stimulation, or if the strength of muscle contraction is more equal in the two eyes in peripheral than in foveal stimulation, we might there have an explanation for the difference in frequency of error in utrocular discrimination under the two stimulation conditions. This is, however, an unverified hypothesis. A possible source of error that

must be considered, but which we think we eliminated, is a difference in an *S*'s ability with foveal and with peripheral stimulation to perceive small displacements of the light on the retina as this results from failure to obtain perfect fixation. Any slight deviation of the light from the center of the fovea is more easily noticed than is a deviation from any other prescribed point. Such a displacement, if it occurred more often temporally or nasally, would afford a position cue to discriminating which eye is stimulated. This cue would be more effective in foveal stimulation. The only way to exclude this source of error is to depend upon the *S* to recognize any displacement and to respond only when stimulation, intended to be foveal, seems to him to be precisely foveal.

Experiment 5.—Results: Utrocular discrimination is better when peripheral stimulation is in the vertical axis than when it is in the horizontal axis.

The same apparatus as in the preceding experiment was used, but the *S*s fixated from trial to trial in clockwise rotation the cross lines above, to the right of, below, and to the left of the stimulation point. These fixation points were .38 in. from the stimulus point. Distance from the eye was 11 in. and convergence was 10 degrees. Each series of 50 trials began with the fixation point next in order to the initial one of the last series. The right-eye and left-eye stimulations were randomized and unpredictable.

Table I shows the errors at each of the four stimulation positions.

TABLE I

PROPORTION OF ERROR IN JUDGING WHICH EYE IS PERIPHERALLY STIMULATED IN THE HORIZONTAL AND VERTICAL AXES

Direction of Stimulus from Fixation Point		Left	Above	Right	Below	Total Horizontal	Total Vertical
Subject	Trials	Error	Error	Error	Error	Error	Error
C	2000	.42	.11	.55	.11	.49	.11
D	400	.27	.09	.30	.10	.29	.09
Total	2400	.40	.11	.51	.11	.45	.11

There was 45 percent of errors in the horizontal meridian and 11 percent of errors in the vertical meridian among 1200 trials in each. This difference is 20 times the S.D. of the difference. There is considerable practice effect, which is shown for the case of Subject C in Table II.

There are two factors which interfere with correct judgments and which cannot be fully controlled in this experiment. One of these is an *S*'s tendency to say 'right eye' more frequently than 'left eye.' The former response was given in 54 percent of the 2400 trials. This, however, occurred as frequently in the vertical as in the horizontal

TABLE II

PRACTICE EFFECT. SUBJECT C. ERRORS IN 100 TRIALS IN HORIZONTAL AND 100 TRIALS IN VERTICAL AXIS EACH DAY

Days	1	2	3	4	5	6	7	8	9	10
Horizontal	64	57	54	48	45	51	49	40	41	37
Vertical	14	17	24	15	15	13	4	8	0	0

meridian. The other interfering factor is the tendency to say 'right eye' when the stimulation is to the right of the fixation point and to say 'left eye' when the stimulation is to the left. This occurred in 53 percent of the 1200 trials in the horizontal meridian. It did not, of course, operate in the vertical meridian, but this slight advantage to the vertical meridian judgments in no way accounts for there being only a fourth as many errors in that meridian as in the horizontal.

Hermans (11) has shown that there is a progressive cyclophoria with increasing convergence and with increasing elevation of the lines of regard. This consists of a rotation about the visual axes by which the upper part of each eye is displaced temporally. This introduces the possibility of a constant stimulation error. With cyclophoria present there would be doubling of peripheral images when there is fusion of foveal images. This means that peripheral stimulation in the vertical axis might seem to the *S* to be displaced to the left in the right eye and to the right in the left eye when the fixation point is above the stimulation point. The reverse of this would occur when the fixation point is below the stimulation point. This factor if present would tend to mislead the *S* when fixation is above and to assist him when fixation is below. It seems clear, however, that this factor did not operate to invalidate the results, as the frequency of error was not only small but was approximately the same for the two fixations in the vertical axis (Table I).

The fact that there is better utrocular discrimination in vertical than in horizontal axis stimulation can establish the truth of no one explanatory hypothesis. It should be remembered that in the superior and inferior rectus the ipsilateral and contralateral reflexes occur in corresponding muscles but in the laterals they occur in opposing muscles. Perhaps there is a greater time lag in the contralateral reflex of the superior and inferior rectus muscles than of the laterals, or perhaps there is less tonus equality of ipsilateral and contralateral reflexes. Such differences would favor utrocular discrimination for stimulation in the vertical axis. Or perhaps in the vertical axis the oblique muscles furnish additional cues. Or it may be that associated changes in eyelid elevation that accompany vertical but not horizontal eye movements would afford a differential cue by which

one can judge which eye is stimulated. Sufficiently precise eye movement photography might throw light on this.

Experiment 6.—Results: With stimulation in the diagonal axes of the visual field utrocular discrimination is significantly, but only slightly, better than chance. The same apparatus, conditions, and general procedure as in the previous experiment were used, except that the fixation points were on the diagonal axes with reference to the stimulation point. To use the oculist's terms, these were axes 45 and 135. The four cross- line fixation points were .54 in. from the stimulus point, this being the hypotenuse of the .38 in. squares. Subject C of the preceding experiment, who had had 2000 trials practice in responding to stimulation in the vertical and horizontal axes, was used.

In 2000 trials, there being 500 at each of the four diagonal fixation points, the proportion of error in judging which eye was stimulated was .455, where .5 is chance. The difference from chance is 2.8 times the S.D. of the difference. If these results are compared with those in Experiment 1, it will be seen that far better utrocular discrimination is secured when the eye is not anchored to a fixation point in an illuminated visual field. In the dark the eye muscles are more relaxed and consequently more free to make reflex responses to the flash of light. No rival response is aroused by a fixation point stimulus.

Experiment 7.—Results: When a single flash of light stimulates one eye or the other, it may be reliably identified as a right eye or left eye stimulus. When two flashes of light are used, both of which at times stimulate one eye, or one of which at times stimulates one eye and one the other, and when these flashes are symmetrically opposite to each other with reference to the fixation point, the S's judgment as to whether one eye or both eyes are stimulated does not deviate significantly from chance.

The apparatus about to be described was used for Experiments 7 and 8. It consisted of a traditional stereoscope through which the *S* looks at two dark gray discs two cm. in diameter on a white celluloid field. The horizontal distance between the centers of the discs is 4.75 cm. At the center of each disc is a right angle *x* composed of .3 cm. black lines. One disc only is seen by each eye, and the two are binocularly fused along with the *x*'s at the centers. The gray discs are actually holes cut in the white celluloid strip exposing a transparent photographic film on which the *x*'s are printed and beyond this a sheet of ground glass. Beyond the glass is a 35 mm. black leader-strip film held in a metal slot. Beyond this, in turn, is the diffusion box containing a neon lamp that provides the stimulus flash.

The essential feature of this apparatus is the black leader-strip in which small holes are punched that admit the light flash to any point or points on either of the gray discs. Between stimulations the leader-strip is moved along a distance of 16 marginal perforations by one rotation of a toothed metal spool. This sets the stimulus pattern for the next trial. The leader-strip, and hence the exposure series, may be of indefinite length. While in use the leader-strip is not rolled but is allowed to come out of and fall into large boxes at either side.

The die used in punching the holes at desired positions along the strip is a steel block drilled at specified points. The strip is drawn through a snugly fitting slot in the block and held in

position by two pins that fit into any of the marginal perforations. Neat round holes .8 mm. in diameter are punched in the leader-strip by passing a rod of drill steel through the holes in the die. The base board that holds the die is ruled off for the distances through which the film must be drawn for punching holes for right or left eye stimulation or for the next exposure pattern in the series.

The stimulus timer used in this set-up has been described in this Journal (15). It is practically noiseless and is operated by a key which the *S* presses when he is ready for the stimulus. The duration of the flash in Experiments 7 and 8 was .015 sec.

Eleven untrained *S*s were first given 40 trials each to determine their ability to distinguish which eye was stimulated when the stimulus was a single flash .75 cm. from the fixation point occurring for each eye in each of 20 directions (Fig. 1). In these 440 trials for all *S*s the proportion of correct judgments was .64, which deviates from chance by more than six times the chance S.D.

Each of these *S*s was next given 40 trials in which the stimulus was two simultaneous flashes placed diametrically opposite to each other with reference to the fixation point. By analogy with a clock dial these might be described as corresponding to 6 and 12 or 3 and 9 o'clock. There were, however, 20, not 12, positions. In one-fourth of the trials both flashes were presented to the right eye, in one-fourth both were presented to the left eye, and in half the trials one of the diametrically opposite flashes was presented to one eye and one to the other. The *S*s were asked to judge whether one or both eyes had been stimulated. In these 440 trials the proportion of correct judgments was .523 which deviates from chance by less than one times its S.D. In both series the order of stimulus position and the order of right, left, or both eye stimulation was randomized and unpredictable.

The use of a stimulus pattern in which the flashes were symmetrically opposite with respect to the fixation point was an attempt, of course, to balance and thus pervert or render ambiguous the specific kinesthetic cues upon which utrocular discrimination may depend. Such stimulation insures the contraction of opposed muscles in both eyes through ipsilateral and contralateral reflex action. The resulting masking of kinesthetic cues would seem to be responsible for the *S*'s failure to discriminate between one-eye and two-eye stimulation.

Experiment 8.—Results: When an irregular four-spot flash pattern is presented so that any two of the flashes stimulate one eye and the other two stimulate the other eye, the S's judgments as to which eye has been stimulated by any specified flash do not deviate from chance.

A factor that seemed likely to invalidate the results in the use of the stimulus pattern in Experiment 7 was the tendency of any well practiced *S* to give more attention to a stimulus point that fell in one-half of the binocular field than he gave to the stimulus point that fell in the other half. If, for example, he habitually came to attend more

to a light in the upper half of the field, then the use of two lights placed in opposite directions from the fixation point might not serve the purpose of balancing the external muscle innervation. The *S* with such directed attention might well move the stimulated eye in the direction of the upper light and so receive a kinesthetic cue for judging which eye was stimulated. The writer while serving as *S* found that such an attention habit was, though unintentional, an easy one to form and a difficult one to overcome, and that it definitely increased the proportion of correct judgments.

In order to render useless this tendency to give attention to only a part of the visual field a new pattern of stimulation was employed in the present experiment.

Four points of light were used to form an irregular pattern in the binocular field. Any two of these lights in unpredictable sequence stimulated the right retina and the other two the left retina. One of the four lights, regardless of the arrangement of the pattern, appeared at each of four equally spaced levels. From top to bottom the lights could therefore be referred to as 1, 2, 3 and 4. Immediately after the pattern of lights had flashed, the *E* spoke a number (1, 2, 3 or 4) and the *S* responded by saying 'right-eye' or 'left-eye' according to whether he thought the right or left eye had been stimulated by that one of the four lights that was named by the *E*.

Because any of the lights might be the one to be identified as a right or left eye stimulus, the *S*'s attention had to be distributed over the entire pattern, and because there were 48 different position combinations of two right eye and two left eye stimulus points, and because the various lights to be identified were called for in unpredictable order, the *S* could form no selective attention habits that would circumvent the purpose of the experiment. This series of 48 patterns was run through four times to form a total series of 192 exposures in order that each dot in each pattern might be designated once as the dot to be judged as a right-eye or a left-eye stimulus.

Twenty *S*s were each given the series of 192 trials. They failed to show any evidence of discrimination. Of the 3840 judgments, .5019 were correct, .5000 being chance. This deviation from chance expectation is one-fourth S.D. of chance variability. *P* is .60.

The .015 sec. exposure preceded by not more than half a second the designation of which dot in the pattern of four was to be judged as a right-eye or left-eye stimulus. The reason for stimulating before designating the point to be judged was that when this order was reversed the *S* tended involuntarily to give selective attention to one point. The purpose here was to make ambiguous the kinesthetic cues, and this was defeated when selective attention was given. Discrimination was rendered more difficult by thus designating the

point to be judged after the stimulus had ceased, but this would not have rendered discrimination impossible if retinal cues alone could serve for utrocular discrimination. This seems to be clearly borne out by the fact that if the right eye flashes are always red and the left eye flashes are green, and if the *S* knows this, utrocular discrimination is 100 percent correct. The *S* has no trouble in remembering the color of any subsequently designated flash. The same is true if cylindrical lenses are placed before the eyes, one at axis 90 and the other at axis 180, so that the flashes are blurred vertically in one eye and horizontally in the other. The memory of the pattern fully suffices for correct responses. But when the difference in the flashes that form the pattern is only that some stimulate the right retina and some stimulate the left retina, it is impossible to discriminate them under conditions that mask fully the kinesthetic cues.

Summary

A light stimulus is not a sufficient sensory cue for discriminating which eye is stimulated by the light. This discrimination depends upon the difference in muscular responses of the eye stimulated and the eye not stimulated. This difference in eye muscle responses furnishes the specific kinesthetic cue for the final discrimination response. Conclusive evidence for the truth of these statements seems to be given by the results here reported. Under conditions that permit maximum difference between right and left eye reaction pattern, utrocular discrimination is maximum (Expt. 1). Under conditions which mask or reduce the kinesthetic pattern difference, discrimination is reduced (Expts. 3, 4, 5 and 6). Under conditions which minimize or eliminate the kinesthetic pattern difference, discrimination approximates zero (Expts. 7 and 8). The internal sphincters play no necessary part in furnishing the specific stimulus cues (Expts. 2 and 7). Differential stimulus patterns produced by the movements of the external eye muscles are the sufficient and probably the necessary specific conditioners for the utrocular discrimination response.

(Manuscript received June 12, 1944)

References

1. Armstrong, H. C. Intensity and temporal factors in utrocular discrimination. (Unpublished thesis.) University of Washington, 1931.
2. Bourdon, B. La distinction locale des sensations correspondantes des deux yeux. *Bull. soc. scient. med. de l'ouest.* Neuvième année, 1900, **IX**, No. 1. Rennes.
3. Bourdon, B. Sur la distinction des sensations des deux yeux. *Année psychol.*, 1903, **IX**, 41.
4. Brückner, A., & v. Brücke, E. Th. Zur Frage der Unterscheidbarkeit rechts und linksäugiger Gesichtseindrücke. *Pflüg. Arch. ges. Physiol.*, 1902, **XC**, S. 290

5. Brückner, A., & v. Brücke, E. Th. Über ein scheinbares Organgefühl des Auges. *Pflüg. Arch. ges. Physiol.*, 1902, **XCI**, S. 360.
6. Brückner, A., & v. Brücke, E. Th. Nochmals zur Frage der Unterscheidbarkeit rechts- und linksäugiger Eindrucke. *Pflüg. Arch. ges. Physiol.*, 1905, **CVII**, S. 263.
7. Fleischl, M. Physiologisch-optische Notizen. I. *Sitzungsber. Wien. Ak. Wiss.*, 3. Agt., 1881, **LXXXIII**, S. 199.
8. Heine, H. Die Unterscheidbarkeit rechtsäugiger und linksäugiger Wahrnehmungen und deren Bedeutung für das körperliche Sehen. *Klin. Monatsbl. Augenheilk.*, 39. Jahrg., 1901, II, S. 615.
9. Heine, H. Zur Frage der Unterscheidbarkeit rechts- und linksäugiger Gesichtseindrücke. *Pflüg. Arch. ges. Physiol.*, 1904, **CI**, S. 67.
10. Helmholtz, H. *Handbook of physiological optics.* (From 3rd German ed.) Optical Soc. of America, 1925. (Part 3 of 1st ed. appeared in 1866.)
11. Hermans, T. G. Torsion in persons with no known eye defect. *J. exp. Psychol.*, 1943, **32**, 307–324.
12. Hofman, F. B. *Die Lehre vom Raumsinn des Auges.* Berlin: Springer, 1920.
13. Peterson, J. Local signs as orientation tendencies. *Psychol. Rev.*, 1926, **33**, 218–236.
14. Schön, W. Zur Lehre vom binokularen Sehen. *Graefes Arch. Ophth.*, 1878, **XXIV**, S. 47.
15. Smith, S. A stimulus timer operating without commutator control. *J. exp. Psychol.*, 1941, **28**, 370–371.
16. Smith, S., & Guthrie, E. R. *General psychology in terms of behavior.* N. Y.: Appleton, 1922.
17. Thelin, E., & Altman, E. R. Identification of monocular functions. *J. exp. Psychol.*, 1929, I , 79–87.

THE ACQUISITION OF A TRACE CONDITIONED RESPONSE AS A FUNCTION OF THE MAGNITUDE OF THE STIMULUS TRACE *

BY BRADLEY REYNOLDS

Northwestern University

Introduction

Few studies have been carried out with the trace conditioned response in American laboratories. This neglect is all the more interesting in view of the fact that it has been accorded no small amount of attention in theoretical writings (**4**, **5**). Present developments in behavioral theories are such as to indicate that even greater utilization of the concept of the stimulus trace and the empirical laws of trace conditioned responses will be made in the future.

The researches the results of which will be reported in this and succeeding papers have aimed at the testing of certain implications of a theory of trace conditioning.

Theoretical Background

The theory presented here consists essentially of a modification of Hull's most recent formulation (**5**). The most convenient procedure for presentation of the theory would seem to be to state the postulate from Hull which deals with the relation of the concept of stimulus trace to learning; and next to give the mathematical statement of the postulate, simplify this statement for our present purposes and then transform the equation to fit the new assumptions which will be made. Finally, consideration will be given to the experimental implications of the modified theory and the experimental data relevant to them.

Hull has formulated his principal assumptions about learning and conditioning into a single postulate as follows: [1]

"Whenever an effector activity (r—R) and a receptor activity (S—s) occur in close temporal contiguity ($\acute{s}Cr$) and this $\acute{s}Cr$ is closely and consistently associated with the diminution of a need ($\dot{G}$), there will result an increment to a tendency ($\Delta_S H_R$) for that afferent impulse on later

* This article is a part of a dissertation presented to the faculty of the Department of Psychology of the State University of Iowa in partial fulfillment of the requirements for the Ph.D. degree. The author wishes to express thanks to Professor Kenneth W. Spence for his advice and assistance.

[1] For a complete discussion of the symbolic constructs used in this postulate and the equation to follow, see Hull (**5**).

occasions to evoke that reaction. The increments from successive reinforcements summate in a manner which yields a combined habit strength ($_{S}H_{R}$) which is a simple positive growth function of the number of reinforcements (N). The upper limit (M) of this curve of habit growth is the product of (1) a positive growth function of the magnitude of need reduction which is involved in primary, or which is associated with secondary, reinforcement; (2) a negative function of the delay (t) in reinforcement; and (3) (*a*) a negative growth function of the degree of asynchronism (t') of $\dot{S}$ and R when both are of brief duration, or (*b*) in case the action of $\dot{S}$ is prolonged so as to overlap the beginning of R, a negative growth function of the duration (t'') of the continuous action of $\dot{S}$ on the receptor when R begins" (Hull, 5, p. 178).

The mathematical formulation is as follows:

$$_{S}H_{R} = M(1 - e^{-kG})e^{-jt}e^{-ut'}(1 - e^{-iN}) \qquad (1)$$

where

M = limit of habit strength under optimum conditions of learning, assigned a value of 100,
e = the base of common logarithms,
G = a constant change in a measurable objective criterion which results in a need reduction (i.e., size of goal object),
t = delay in reinforcement,
t' = $T_R - T'_{\dot{S}} - .44$ where $\dot{S}$ and R are of brief duration,
T_R = the time of beginning of R,
$T'_{\dot{S}}$ = the time of beginning of $\dot{S}$,
N = the number of reinforcements,
k, j, u and i = empirical constants.

In accord with Hull's formulation it will be assumed that the limit of the curve of habit strength is a function of the time of delay of reinforcement (t) and the size of the goal object (G). The modification to be suggested is concerned with the third variable—the conditioned stimulus and its temporal relations to the unconditioned stimulus (or response). For purposes of simplification we may define the maximum value of $_{S}H_{R}$ for a given constant set of values for t and G as M'. Thus:

$$M' = M(1 - e^{-kG})e^{-jt}. \qquad (2)$$

The new assumption to be made is that the limit of habit strength will be a function also of the magnitude of the stimulus trace (s) and the maximum value (M') as determined by given constant values of t and G. Employing the symbol M'' to represent this value, the assumption may be formulated as follows:

$$M'' = M' \times s. \qquad (3)$$

It remains to specify the experimental variables that determine the value of the stimulus trace. The assumption to be made is that the magnitude of stimulus trace is a function of three variables. The

first is the interval between the conditioned and unconditioned stimulus. It is held that as the interval between the conditioned stimulus increases from 0 to .45 sec. the magnitude of the stimulus trace increases as a positive function of the interval. As the interval between conditioned and unconditioned stimulus increases beyond .45 sec. the magnitude of the stimulus trace declines as a negative growth function of the interval.[2]

The second variable which enters into the determination of s is the ordinal position of a given presentation of the conditioned stimulus in a series of presentations. As the ordinal position of the presentation of the conditioned stimulus increases numerically, s decreases as a simple decay function of the ordinal position of the conditioned stimulus, in a series of such presentations.

Finally, it is postulated that the rate of presentation of the stimulus (S) determines the lower limit which the magnitude of the stimulus trace will approach with the ordinal number of the presentation of the stimulus. The complete formulation of these postulates may be stated as follows:

$$s = Pe^{-[F(1-e^{-jN'})+kt']} \tag{4}$$

where

F = the rate of presentation of the conditioned stimulus,
N' = $N - 1$, where N is the number of presentations of the conditioned stimulus,
t' = $T_{us} - (T_{cs} - .45)$,
T_{us} = the beginning of the unconditioned stimulus,
T_{cs} = the cessation of the conditioned stimulus when it is of brief duration,
k, j and P are constants.

Substitution for s in equation 2 gives

$$M'' = M'Pe^{-[F(1-e^{-jN'})+kt']}.$$

Since M' and P are constants we may conveniently substitute a third constant (M^0) which represents the product of M' and P. Then

$$M'' = M^0e^{-[F(1-e^{-jN'})+kt']}.$$

and, rewriting the equation 1,

$${}_sH_R = M^0e^{-[F(1-e^{-jN'})+kt']}(1 - e^{iN}).$$

[2] It should be noted that the optimal interval between conditioned and unconditioned stimulus as given here differs from that given by Hull by .01 sec. Hull arrives at his value of .44 sec. from a mathematical treatment of Wolfle's data (8). We have taken the value at .45 sec. purely for the sake of convenience. There is general agreement among experimenters in the field of conditioning that the optimal interval between conditioned and unconditioned stimulus is somewhere in the neighborhood of .4 to .5 sec. There hardly seems to be sufficient evidence at present to set the value exactly. The value used here, .45 sec., lies midway between .4 and .5 sec.

If we now assume with Hull that habit strength ($_sH_R$) is one determinant of the strength of the conditioned response (e.g., the percentage of CR's in a block of trials) we can draw certain implications as to the limits of conditioning under variations of the experimental conditions referred to in the above theoretical formulation. We shall consider two implications of this theory.

The first implication regards the relation of the limit (M) and rate of rise of the acquisition curve for the trace conditioned response to the measure (t') based upon the length of the interval between the conditioned and unconditioned stimulus. If the frequency of presentation (F) is held constant, then the limit (M) and the rate of rise of the learning curve will decrease as a simple decay function as the interval between the conditioned and unconditioned stimulus increases beyond .45 sec., or decreases from .45 sec. Stated more specifically, the limit (M) and rate of rise of the curve of acquisition for a trace conditioned response will be less when the interval between the conditioned and unconditioned stimulus is .35 sec. than when it is .4 or .45 sec. In the same manner the limit (M) and rate of rise of the learning curve will be less when the interval between the conditioned and unconditioned stimulus is .75 sec. than when it is .65 or .45 sec. This implication may be made more apparent by the following figure (Fig. 1).

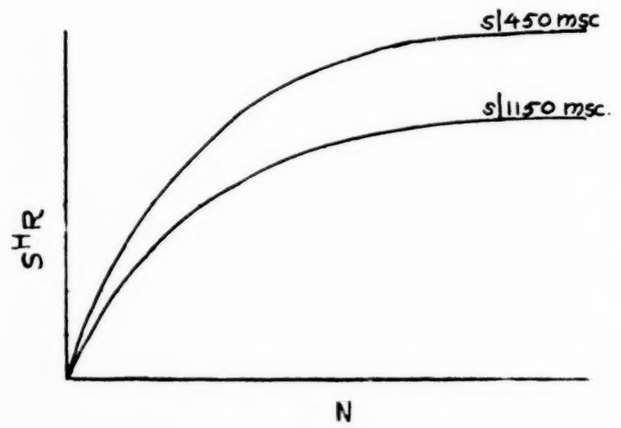

FIG. 1. Theoretical growth of habit strength with a stimulus trace of 450 msc. and a stimulus trace of 1150 msc. when the rate of presentation of the conditioned stimulus is held constant.

This figure shows the theoretical curve of growth of $_sH_R$ for a stimulus trace of 450 msc. and a stimulus trace of 1150 msc. when F is held constant. That is to say, if we are presenting our conditioned stimulus once per min., then the rate of rise and the limit (M) of the acquisition curve for the conditioned response for a stimulus trace of 450 msc. will be greater than that observed in the acquisition curve for a response conditioned to a stimulus trace of 1150 msc. The

part of this implication which relates to the limit (M) of the acquisition curve may be illustrated in another way. This involves plotting theoretical limits for acquisition curves against the measure (t') based on the interval between the conditioned and unconditioned stimulus. This has been done in the figure below (Fig. 2).

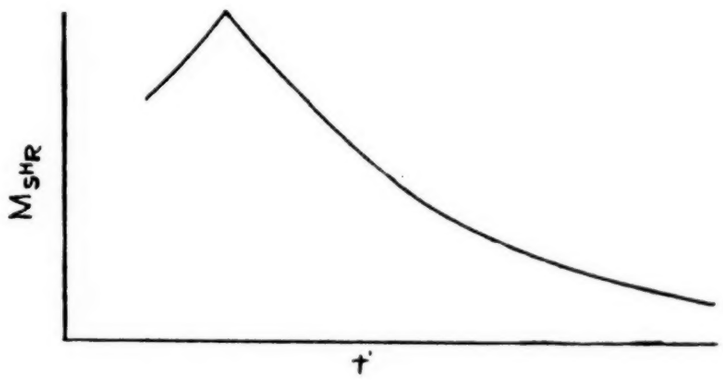

FIG. 2. Theoretical limits of the growth of habit strength as a function of the interval between conditioned and unconditioned stimulus.

The second implication involves the relation of variations in the rate of presentation of the conditioned stimulus with changes in the limits and rates of rise of the acquisition curves for trace conditioned responses. If t' is held constant, then as F increases, the limit (M) and rate of rise of the acquisition curve will decrease. If we are dealing, for example, with a stimulus trace of 450 msc., the acquisition curve for a response conditioned to this stimulus trace will rise more rapidly and have a higher limit if the conditioned stimulus is presented once per min. than if it is presented twice per min. The relationship is identical with that illustrated above in Fig. 1 except that we are concerned with the same stimulus trace and two different rates of presentation.

There is some positive confirmation for the first implication drawn above, regarding limits (M) of acquisition curves and their relation to the length of interval between the conditioned and unconditioned stimulus. This is found in a study of Wolfle (**8**), the first of two such studies carried out by this experimenter with the trace conditioned response. The conditioning technique used was finger withdrawal to shock. Nine groups of 10 Ss, one group for each of nine different intervals between conditioned and unconditioned stimulus, were given 340 paired presentations of conditioned and unconditioned stimulus, and 60 test trials with conditioned stimulus alone. Hull (**5**) has fitted an exponential function of the type plotted in Fig. 2 above to the data from this experiment, and it will be observed that the fit is very close. The optimal interval between

conditioned and unconditioned stimulus was .5 sec. These results are in essential agreement with Hull's theory and the theory presented here. Wolfle does not present learning curves, however, so that confirmation as to limits and rates of rise of acquisition curves cannot be determined from these data. A further difficulty is that the measure of conditioning used is percent response in 60 test trials. The acquisition curves for two conditioning groups may subtend identical areas under the curves and still the curves may have different limits.

Wolfle has reported a second study with the trace conditioned response using the same technique of conditioning as used in the study noted above (**9**). Twelve groups—one group for each of 12 different intervals between conditioned and unconditioned stimulus—with from five to seven *Ss* in each, were given from 600 to 1200 paired presentations of the conditioned and unconditioned stimulus. In this study the greatest percentage of conditioned responses occurred when the interval between conditioned and unconditioned stimulus was .2 sec. This interval is shorter than the optimal value postulated in the present theory. This inconsistency between the results of the two studies by Wolfle will be discussed later.

There is a third study, that of Bernstein (**1**), which studied the relation of the temporal interval between conditioned and unconditioned stimulus and frequency of conditioned response during the training trials. The response used was eyelid closure to shock. The optimal interval observed in this study was .3 sec. but no significant differences were observed between any interval between .3 sec. and 1.48 sec. This is contradictory for both Hull's theory and the present theory. It should be noted, however, that only 50 paired presentations of conditioned and unconditioned stimulus were given in this experiment. The relevance of this factor of limited reinforcements in the determination of such results as Bernstein reports will be considered in the discussion of the present investigation.

The implication drawn from the theory regarding the distribution of trials is consistent with such relevant data as exist. Calvin (**3**), in what is apparently the only systematic study concerned with the distribution of trials, found that increasing the rate of presentation of paired stimuli from 3 to 18 per min. decreased the rate of rise and the limit of acquisition curves for the conditioned eyelid response. Humphreys (**6**) reports a study with the eyelid response in which conditioning was identical for groups to which 48 trials were given per session for two sessions, and groups which received half as many trials in the same period of time. The two inter-trial intervals were 30 and 60 sec. Schlosberg (**7**) in a study with conditioned breathing responses in the white rat, found it less advantageous

to use 200 paired presentations of conditioned and unconditioned stimulus per day than to use 25 paired stimuli in the same period. There is no direct evidence relevant to the implication for the relation of distribution of trials to the acquisition of a trace conditioned response.

Purpose

The purpose of the present investigation was to test the implications drawn from the theory of trace conditioning discussed in the preceding section. We may summarize these briefly again.

1. *The relation of interval between conditioned and unconditioned stimulus and the rate of rise and limit of the acquisition curve for a trace conditioned response.*

The implication of the theory is that with increasing length of interval (with the exception of very short intervals as noted) the acquisition curve will rise more slowly and approach a lower maximum value. In order to test this implication, *S*s were conditioned with four different intervals between the conditioned and unconditioned stimulus. The four intervals were chosen so as to have intervals equal to, greater than, and less than the postulated optimal interval.

2. *The relation of the rate of application of the conditioned stimulus to the rate of rise and the limit of the acquisition curve for a trace conditioned response.*

For purposes of testing this implication, the four interval groups were split, and half of the *S*s were trained with massed trials and half with distributed trials. The implication to be drawn from the theory is that with massing the rate of rise and the limit of the acquisition curve would be less than with distribution of the training trials.

Procedure

Apparatus

The apparatus used in the present study was the Dodge photochronograph which has been described elsewhere (2). Recording of the eyelid movement was by the paper eyelash method. The unconditioned stimulus was an airpuff delivered by the fall of an 80 mm. column of mercury in a manometer. The air tube which delivered the puff to the *S*'s eye divided and activated a paper reed which recorded a shadow on the photographic record. Calibration of the moment of the incidence of the puff at the *S*'s eye and the recorded movement of the paper reed showed no difference in time between the two, within the limits of the recording system used. The conditioned stimulus used was a click delivered to the *S* through a pair of earphones. This click was taken from an electromagnetic signal marker through a paper, tubular condenser. The signal marker carried a bamboo reed which interrupted the beam of light when this click was delivered. The incidence of the click and the recording of it were necessarily simultaneous. The signal marker was mounted in back of the lens on the lensboard and was unshielded. The sound of its movement combined with the click, delivered through the phones, to form the stimulus. There was an on and off phase of the click, the whole stimulus event consuming 50 msc.

Procedure

A total of 77 *Ss* were used in the experiment. They were divided into two groups, one of which was trained with massed trials and one with distributed trials. Each of these groups was further divided into four subgroups, each of which was conditioned with a different interval between the conditioned and unconditioned stimulus.

The distributed group.—Thi . group was divided into four subgroups, one for each of the following intervals between conditioned and unconditioned stimulus: 250 msc., 450 msc., 1150 msc., 2250 msc. These *Ss* were given 90 reinforcements with intervals between the trials of 1–2 minutes. *Ss* were trained through four sessions, 25 reinforced trials being given in the first three sessions and 15 reinforced trials in the last. Whenever the interval between trials was two min., the *S* was allowed to remove his head from the head stand and rest for that period.

The massed group.—This group was divided in the same manner as the distributed group and conditioned with the same intervals between conditioned and unconditioned stimulus except for the longest interval. In working with this interval with the distributed group, certain mechanical difficulties which tended to slow up the rate with which trials could be run, were encountered. This created no inconvenience with distributed trials, but would have seriously hampered the giving of massed trials. By making some adjustments with the pendulum release mechanism and shortening the interval somewhat, these difficulties were overcome. The longest interval between conditioned and unconditioned stimulus for the massed group was thus 2100 msc.

Ss were given the same number of reinforced trials as the distributed group, but two sessions instead of four were used. In the first session 60 reinforced trials were given, and 30 reinforced trials were given in the second session. For the 2100 msc. group only one session was used, 90 reinforced trials being given with no extinction trials.

The interval between trials was 10–20 sec. with one two-min. rest period in each session.

All *Ss* were given five adaptation trials with the conditioned and unconditioned stimulus alone prior to training with paired stimuli. The instructions to the *Ss* were meant to have them assume a neutral set. They were asked not to question *E* as to what the experiment was about, nor to try to guess the purpose of the experiment, but to make an effort to put the experiment out of mind if they found themselves thinking about it. In the main the *Ss* were undergraduate women from the first course in psychology at the State University of Iowa, and all but some 15 were volunteers.

Results

Acquisition

The distributed group.—The acquisition curves for the four interval groups in this main group have been plotted and are presented in Fig. 3. Each point on the curves represents the mean frequency for a block of 10 trials expressed as a percent. For most curves this percent is equivalent to the total number of responses in the block, since the number of *Ss* is 10. An examination of the curves reveals that, with the possible exception of the curve for the 250 msc. group, the acquisition curves have reached a limit. The limit for the 450 msc. group is near the maximum possible percent of response. Below the limit for the 450 msc. group are in order of size the limits of the 1150 msc., 250 msc. and 2250 msc. groups.

When the rates of rise are considered it can be seen that the initial rate of rise of the 1150 msc. group is greatest, with that of the 2250 msc. group next, those of the 450 msc. and 250 msc. groups following in that order. When the rate of rise is determined from the 1–10th trial point the rates of rise of the curves are in the same order as the limits approached by the respective curves. In the dis-

cussion of the results, an explanation will be given for the change in the relation of the different rates of rise with increasing trials.

The observed regularity of the curves can be attributed to two factors. First, each point is determined upon the basis of 100 trials. For certain theoretical purposes this method of plotting curves has certain disadvantages, although the method is that commonly used for plotting such data. For our present purposes the method offers little if any disadvantage; even a grosser unit than trials by tens could have been employed. A second factor which undoubtedly

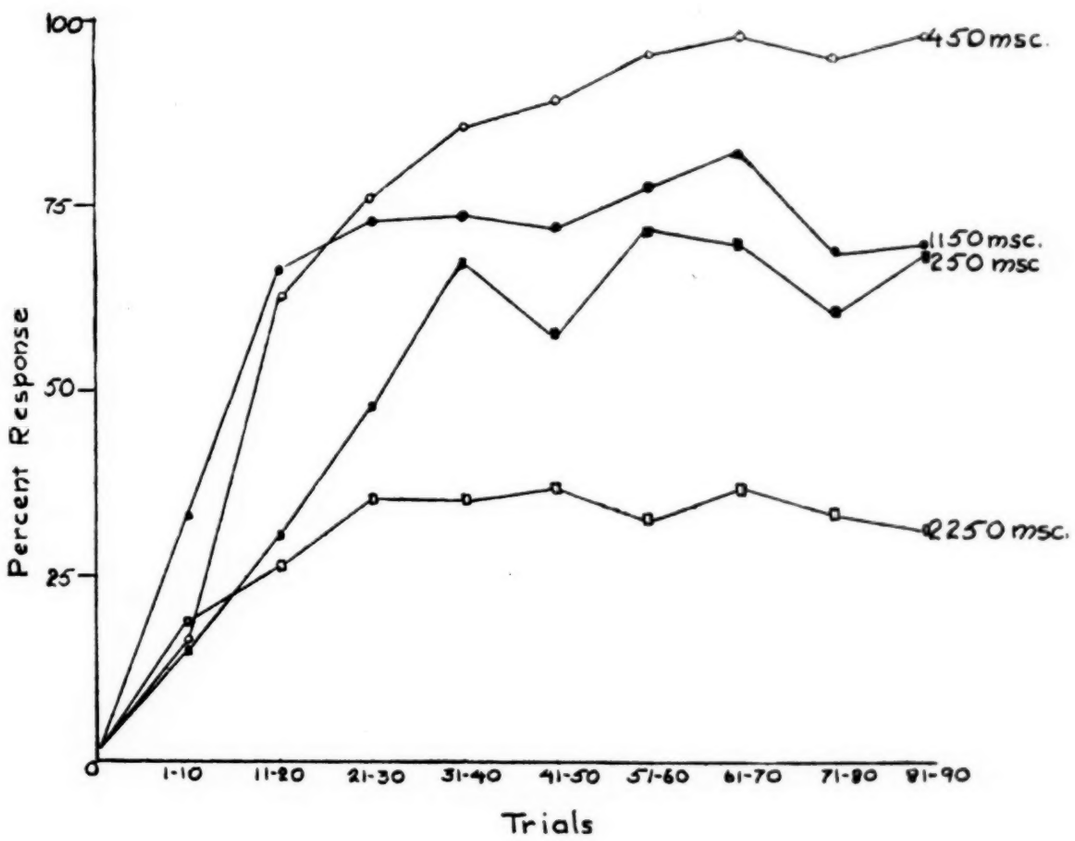

FIG. 3. Acquisition curves for trace conditioned responses when the mean rate of presentation of the conditioned stimulus is one per one and one-half min. (i.e., under conditions of relative distributing of the acquisition trials).

contributed to the regularity of the curves is the size of the interval between trials, 1–2 min., and the spacing of the 90 reinforcements through four days. No study with the eyelid response has been reported in which distribution and spacing of trials as great as this has ever been employed. The usual procedure has been to present trials as rapidly as the apparatus would permit, giving 100 trials in one or two sessions. Total responses for the complete training series expressed as percents for each S are presented in Table III. The means and sigmas of these percents are given for each group. These last

have been calculated directly rather than from the grouped data as presented here.

The massed group.—Data for the massed group have been presented in the same fashion as for the distributed group (Fig. 4, Tables II and III). An examination of the acquisition curves reveals that all the curves, with the exception of that for the 450 msc. group, have apparently reached a limit. The 450 msc. group is approaching the highest limit, the other groups in order of their limits are the 250 msc., the 1150 msc. and the 2100 msc. groups.

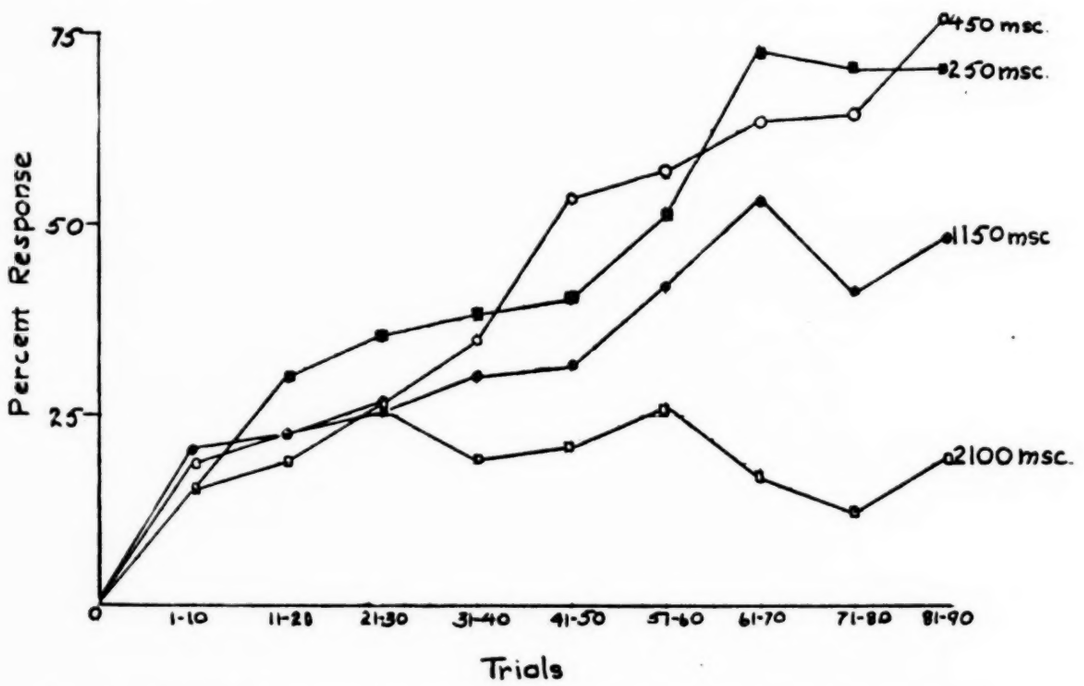

FIG. 4. Acquisition curves for trace conditioned responses when the mean rate of presentation of the conditioned stimulus is four per min. (i.e., under conditions of relative massing of the acquisition trials).

TABLE I

PERCENT OF CONDITIONED RESPONSES FOR SUCCESSIVE 10 TRIALS FOR EACH INTERVAL GROUP WHEN TRAINING TRIALS WERE DISTRIBUTED

Trials	Percent Conditioned Response			
	250 msc.	450 msc.	1150 msc.	2250 msc.
1–10	15	16	33.3	19
11–20	30	62	66.7	26
21–30	48	76	73.3	35
31–40	67	85	74.4	35
41–50	57	89	72.2	37
51–60	72	96	77.8	32
61–70	70	98	82.2	37
71–80	60	95	68.9	33
81–90	68	98	70.0	31
N =	10	10	9	10

TABLE II

PERCENT OF CONDITIONED RESPONSES FOR SUCCESSIVE 10 TRIALS FOR EACH INTERVAL GROUP WHEN TRAINING TRIALS WERE MASSED

Trials	Percent Conditioned Response			
	250 msc.	450 msc.	1150 msc.	2250 msc.
1–10	15	18.9	21.1	15
11–20	29	22.2	22.2	19
21–30	35	26.7	25.5	26
31–40	38	34.4	30.0	19
41–50	40	53.3	31.1	21
51–60	51	46.7	42.2	26
61–70	73	63.3	53.3	16
71–80	71	64.4	41.1	12
81–90	71	77.8	48.9	20
N =	10	9	9	10

TABLE III

PERCENT OF CONDITIONED RESPONSES FOR EACH *S* FOR THE TOTAL TRAINING SERIES

Percent CR	Distributed				Massed			
	250 msc.	450 msc.	1150 msc.	2250 msc.	250 msc.	450 msc.	1150 msc.	2100 msc.
90–99		2	3			1		
80–89	1	3	1					
70–79	2	4		1	2		2	
60–69	2	1	2		3			
50–59	1		1	2	1	2		
40–49	1			1		2	2	3
30–39	1		2		1	3		
20–29	2				1	1	3	2
10–19				6	1			1
0–9					1		2	4
M	48.7	73.5	61.9	26.4	43.4	41.9	32.3	17.3
σ	18.7	9.5	18.7	21.2	21.6	16.7	8.6	15.7

The relation of the rates of rise of the several curves is somewhat similar to that holding for the distributed group. The 1150 msc. group shows the fastest initial rise, with the 450 msc. group next, the 250 msc. next and the 2100 msc. group showing the slowest rate of rise. Due to the irregularity of the curves it is hard to determine any regular change in relation of the rates of rise of the curves.

In view of the fact that Bernstein's study with trace conditioning used intervals between trials comparable to those used here, it is interesting to compare the curves early in the trials and late in the trials with his results. It will be remembered that Bernstein found no significant difference for any interval between 300 msc. and 1480

msc. with 50 reinforced trials. If we examine the present curves at the 21–30 trial, we see that there is no difference among the intervals above 250 msc., with the latter point showing a slightly greater percent of responses than the others. However, at the 31–40th trial point it will be observed that the curves are in exactly the same relation as the limits for the curves. The inconsistency between the two studies is thus apparently resolved.

An inconsistency between the results of the two Wolfle studies was noted above. It was pointed out that the optimal interval between conditioned and unconditioned stimulus was .5 sec. in the first study and .2 sec. in the second. Wolfle's conditions of rate of presentation of the conditioned stimulus were approximately those for the massed group in the present study. Wolfle's measure of conditioning was percent response for the total training period as determined by test trials distributed throughout that period. It should be noted that if this measure is employed in the present study, the 250 msc. group shows more conditioning (43.4%) than the 450 msc. group (41.9%), although the latter group reaches a higher limit.

Discussion of Results

The distributed group.—The first implication drawn from the theory states that with *increasing* intervals between conditioned and unconditioned stimulus beyond 450 msc. the limits of curves of acquision will decrease. Similarly, for *decreasing* intervals between conditioned and unconditioned stimulus short of 450 msc. the limits will also decrease. An examination of Fig. 3 shows that this implication is confirmed. The implication further states that the limit of the acquisition curve for a trace conditioned response is a negative growth function of the length of the interval beyond 450 msc. The limits of the acquisition curves for the distributed groups for each interval have been plotted against intervals, and the points joined by straight line segments (Fig. 5). It is hardly feasible to fit as few points as are plotted here to any mathematical function. However, the general shape of the plot can be seen to be consistent with the implication drawn from the theory. The implication also states that the limit of the acquisition curve for a trace conditioned response is a positive function of the length of the interval between conditioned and unconditioned stimulus for intervals less than 450 msc. A consideration of Fig. 5 shows that the relation of the point for the 250 msc. interval to that for 450 msc. is not inconsistent with this implication.

The second implication of the theory is that the rate of rise of the acquisition curve for a trace conditioned response will *decrease* with increasing intervals between conditioned and unconditioned stimulus

beyond 450 msc. Further, for intervals between conditioned and unconditioned stimulus shorter than 450 msc. the rate of rise of the acquisition curve will *increase* with increasing intervals. If we consider the rates of rise of the acquisition curves (Fig. 3) from the 1–10th trial ordinate we can see that the observed rates of rise are in complete agreement with the implication of the theory.

It should be noted that the rate of rise of the acquisition curve for the 450 msc. and 250 msc. interval groups is less than that for the 1150 msc. and the 2250 msc. groups between the 0 and 1–10th trial

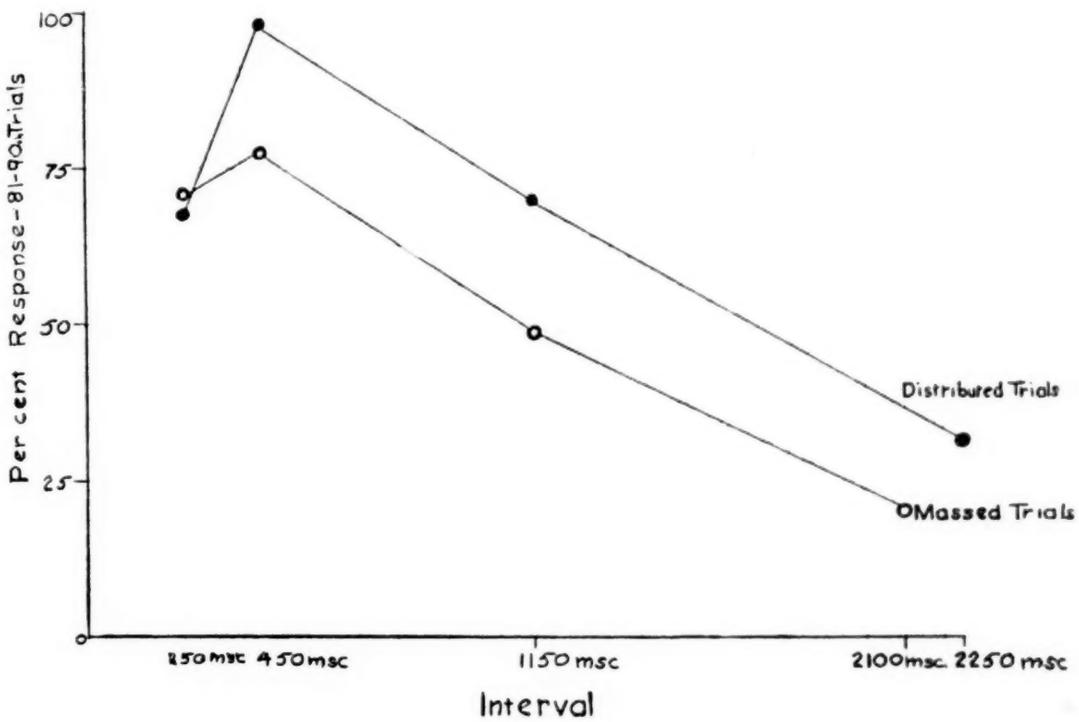

FIG. 5. Limits of acquisition curves for trace conditioned responses under conditions of relative massing and relative distributing of the acquisition trials.

ordinates. The explanation of the deviation from the predicted relationship may be found in the relation of the latencies of conditioned responses early in the training period and the length of the interval between conditioned and unconditioned stimulus. It is known that the latency of the conditioned eyelid response decreases with increasing number of trials (3). When the interval between the conditioned and unconditioned stimulus is very short, conditioned responses whose latencies are longer than the interval cannot be recorded, due to the fact that the unconditioned stimulus occurs before the response can be made. This will be less true in the case when longer intervals between conditioned and unconditioned stimulus are used. To measure all the conditioning present in earlier trials, in the case of shorter intervals, it would be necessary to omit the

unconditioned stimulus, as is done in salivary conditioning with overlapping stimuli.

The massed group.—The implication drawn from the theory regarding rate and limit of the acquisition curve for a trace conditioned response in the case of massing of the trials is identical with that considered above, for relations holding within the massed group. When we examine the acquisition curves (Fig. 4) we can immediately see that this prediction as to limits is confirmed. The interval determines the limit, and a consideration of the plot of limits against intervals (Fig. 5) shows that the relationship is in agreement with the theory.

The implication for rate of rise would also seem to be confirmed by the results. The relationship between the curves for short and long interval groups in the early trials is essentially the same as with the distributed group, with the exception that the rate for the 2100 msc. is lowest in the case of massing. When we compare rates of rise between later intervals, however, we see that the rates of rise are in agreement with the implication of the theory.

Distribution vs. massing in acquisition.—It will be remembered that the implication of the theory states that the acquisition curve for a trace conditioned response when training trials are distributed will rise more rapidly and reach a higher limit than when training trials are massed. Considering either the two sets of acquisition curves (Fig. 3 or 4), or the plot of limits against intervals (Fig. 5), we can see that in the main this holds for limits of the acquisition curves for both groups. It should be noted that the curve for the 450 msc. massed group has not reached a limit. However, if the curve were extended without any marked change in rate the apparent limit is somewhere in the vicinity of 80–84 percent which would be well below that for the distributed group at the same interval. The only real deviation from the predicted relationship is that of the 250 msc. groups, the massed group having a limit slightly above that of the distributed group. The explanation of this has already been suggested. It was pointed out that short intervals between conditioned and unconditioned stimulus would prevent the appearance of long latency responses and that the effect would be more marked in the case of the 250 msc. interval and would operate for a longer period of time. When the latencies of these obscured responses become shortened they will tend to come in with a 'rush,' and a marked rise in the curve will occur. The curve will then rise more slowly to a maximum. This is apparent in the curve for the massed group. On the other hand the curve for the distributed group rose approximately nine percentage points between the 80th and 90th trials. There is more than a strong suggestion that with continued trials the

limit of the curve would have been above 80 percent, which is near to the predicted limit. A more apt comparison between the two groups under these circumstances would be based on percent response for the complete training period of 90 trials. This is approximately 48.7 percent for the distributed group and 43.4 percent for the massed group.

Comparison of differences in rates of rise and limits of acquisition curves for massed and distributed groups can be made without qualification between curves which have reached limits and in which the rate of rise has not been depressed through the masking of long latency responses in the early trials. The 1150 msc. and 2250 and 2100 msc. groups can be assumed to meet these conditions. An examination of the curves for these groups (Fig. 6) reveals that the

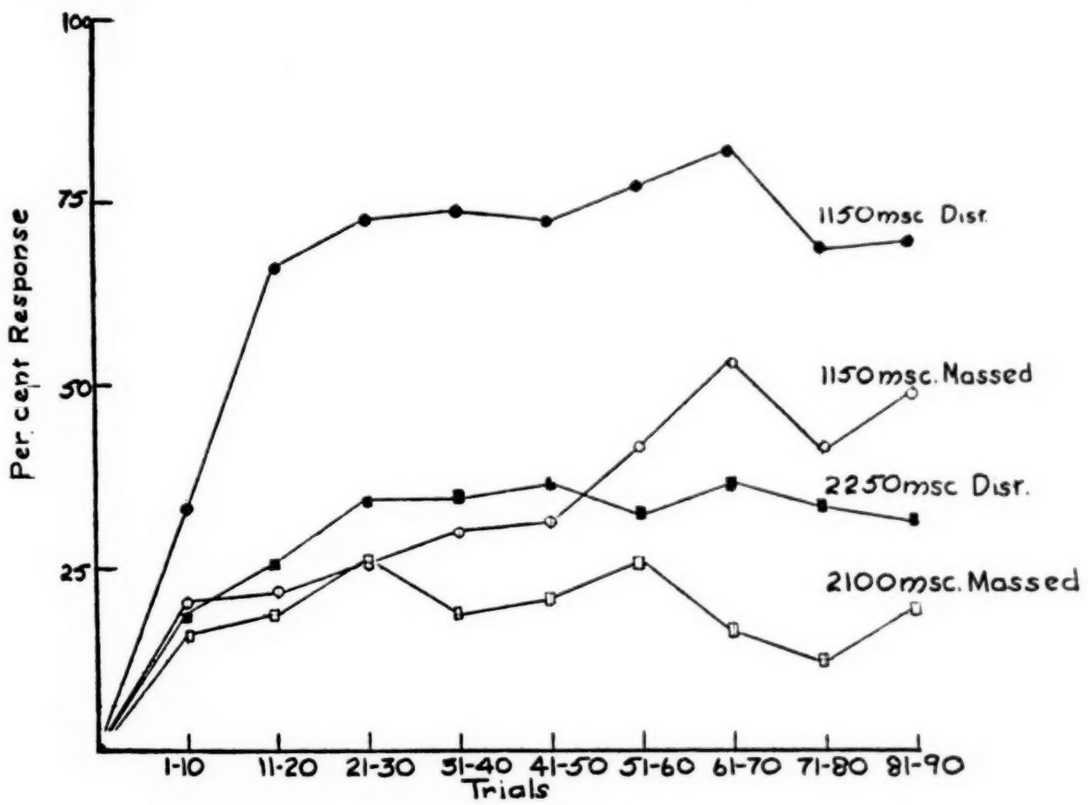

FIG. 6. Acquisition curves for trace conditioned responses with intervals between conditioned and unconditioned stimulus of 1150, 2250 and 2100 msc., under conditions of relative massing and distributing of the acquisition trials.

rate of rise and the limit of each distributed group is greater than that for the corresponding interval group trained with massed trials. This is in agreement with the theory.

SUMMARY

In the present study two sets of relations have been investigated. These relations are:

1. The rate of rise and limit of the acquisition curve for a trace conditioned response as a joint function of the interval between the conditioned and unconditioned stimulus, and the rate of presentation of the conditioned stimulus.

2. The 77 Ss were divided into two groups for conditioning the eyelid response to a stimulus trace. One group, the distributed group, was given 90 reinforced trials with intervals between trials of 1–2 min., this training being spread through four sessions. The second group, the massed group, was trained under the same conditions except that the interval between trials was 10–20 sec., and 90 reinforcements were given in two sessions. Each of the two groups, massed and distributed, was divided into four groups, upon the basis of the interval between the conditioned and unconditioned stimulus. The intervals employed were 250 msc., 450 msc., 1150 msc. and 2250 msc. in the distributed group and 250 msc., 450 msc., 1150 msc. and 2100 msc. in the massed group.

3. The results showed that the limits and rates of rise of the acquisition curves decreased for intervals shorter than or longer than 450 msc., the function approaching in form the exponential type.

4. It was found that when acquisition trials were massed the limits and rates of rise of the acquisition curves for any given interval between conditioned and unconditioned stimulus were, with one exception, less than when acquisition trials were distributed.

(Manuscript received May 26, 1944)

Bibliography

1. Bernstein, A. L. Temporal factors in the formation of conditioned eyelid reactions in human subjects. *J. gen. Psychol.*, 1934, **10**, 173–197.
2. Dodge, R. A pendulum-photochronograph. *J. exp. Psychol.*, 1926, **9**, 155–161.
3. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* New York: D. Appleton-Century, 1940.
4. Hull, C. L., Hovland, C. I., Ross, R. T., Hall M., Perkinds, D. T., & Fitch, F. B. *Mathematico-deductive theory of rote learning.* New Haven: Yale University Press, 1940.
5. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
6. Humphreys, L. G. Distributed practice in the development of the conditioned eyelid reaction. *J. gen. Psychol.*, 1940, **22**, 379–385.
7. Schlosberg, H. An investigation of certain factors related to ease of conditioning. *J. gen. Psychol.*, 1932, **7**, 328–342.
8. Wolfle, H. M. Time factors in conditioning finger-withdrawal. *J. gen. Psychol.*, 1930, **4**, 372–378.
9. Wolfle, H. M. Conditioning as a function of the interval between the conditioned and the original stimulus. *J. gen. Psychol.*, 1932, **7**, 80–103.

TRIAL AND ERROR LEARNING WITH DIFFERENTIAL CUES

BY HAROLD E. JONES[1]

University of California

For three decades, students of the learning process have recognized the significance of problems in 'the fundamental science of modifiability by satisfiers and annoyers.' Although not always described in the Thorndikian idiom (**34**, p. 313), rewards and punishments as related to learning have figured in numerous experimental studies. Among these have been experiments dealing with the effect of amplifying or supplementing the reward for correct responses, and the effect of intensifying the punishment for errors. Results from such investigations have led to inquiries as to the role of emphasis or punishment applied to the right rather than the wrong reactions, and also as to the effect of using accenting or punishing stimuli non-specifically, i.e., independent of particular choices in the learning process.

It must be admitted that the literature in this field has failed to show any very conspicuous unity in conclusions. A representative sampling of studies may be cited as follows:

1. Rewarding right responses is more effective than punishing wrong responses (**22**, **28**, **30**, **35**). Positive guidance (emphasizing right responses) is more effective than negative guidance (emphasizing wrong responses) (**16**, **17**, **27**, **29**, **37**). Negative guidance has little or no effect upon learning (**20**, **32**, **40**). And on the other hand. . . . It is more important to punish wrong responses than to reward right responses (**7**, **15**, **41**). Informative emphasis on incorrect responses (negative guidance) is an aid to learning (**5**, **12**, **13**). Emphasizing wrong responses is about as effective as emphasizing right responses (**8**, **14**, for light shock; **26**, for measures of retention; **25**).

2. Emphasis aids learning even if it is non-informative (e.g., light or shock given during the experiment but not applied to any specific choice, **12**). . . . Non-informative emphasis has no effect upon learning (**4**).

3. The effect of punishment or emphasis is general rather than specific, influencing total performance rather than the part responses

[1] The writer is indebted to Professor Warner Brown for reading the ms. and for a number of helpful suggestions, to Dr. Joseph A. Yoshioka for statistical assistance, and to Mr. Jerome Rulon for assistance in administering the tests.

which it accompanies (**8, 24**). . . . Punishment or emphasis has not merely a general but also a specific effect in 'marking' blind alleys (**1**).

Numerous attempts have been made to draw educational implications from experiments in this field. It would appear, however, that almost any statement that can be made about the role of reward, punishment, and emphasis in learning can be countered with an opposite statement, based on the evidence of one or more experimental studies.

The situation is of course not quite as confused as would appear at first glance. In some instances, conclusions which are in apparent conflict rest upon statistically insignificant differences, or upon experimental comparisons which have failed to make use of comparable groups. Variable effects upon the learning process may also, in some instances, be accounted for in relation to the task, the subjects, the motivational devices, and whether or not the experimental procedure was directed toward modifying specific part-responses, or toward changing attitudes or incentives. Quantitative relationships are important. Thus, when it is shown (as for example by Rock, **28**) that supplementary rewards have little or no effect, this can be considered in relation to findings that although an increase in the reward favors learning (**9, 11, 22, 36**), a point of diminishing returns may be reached beyond which learning is no longer aided (**36**). If, therefore, an experiment is undertaken at a high motivational level, the introduction of an additional reward would be of less importance than at lower motivational levels, and might even become a disturbing factor. Similarly, although an increasing degree of punishment is sometimes reported as favoring learning (**2, 3, 5**), punishment beyond a certain intensity will fail to produce gains and may involve losses (**14, 30, 39**). Differential effects may depend in part on the strength of the connections which are exposed to reward or punishment; thus, Stephens (**31**) reports the greater influence of reward upon weak associations, and the greater influence of punishment upon strong associations. Recitation or inner repetition may also be a factor in determining conflicting results, since it has been shown that punishment is influential if the subject is given the opportunity to rehearse what he has learned, but that without the opportunity for confirming repetition, punishment may be without effect (**32**). Clearly, if conclusions must apply so narrowly to specific areas and conditions, and must be hedged about by so many empirical qualifications, we have not yet reached a stage at which generalizations can be very impressive. It would seem wise, for the time being, to delay attempts to unite these diverse findings into a single theoretical formulation, and to caution against the setting up of a school of educational method, with procedures deduced from a 'law' of learning.

So far as educational implications are concerned, one of the defects of this great mass of material is that except in a few instances, it rests either upon animal studies or upon studies of a highly selected sample of human learners, namely, college undergraduates and graduate students. More evidence is needed, both from laboratory and classroom studies, as to factors in learning among representative groups of school children. The present investigation, dealing with pupils from a public junior high school, is directed toward the problem of the effect upon learning of informative signals, when these are experimentally varied as to their intrinsically satisfying or annoying character. Electric shock, the favorite 'emphasizer' in laboratory experiments, has been avoided, partly because it is so remote from ordinary experience and partly because individuals show such wide and unpredictable variations (**21**, **38**) in their response to shock. Since it seemed desirable to examine the differential role of stimuli where these fall within a more normal range as to satisfaction or annoyance, an attempt has been made to utilize mildly pleasant or preferred, indifferent, and mildly disagreeable or rejected signals, and to apply these in a balanced order to both the correct and incorrect responses in trial and error learning.

Sample

Fifty-seven boys and an equal number of girls participated in the experiment. A random subsample of a larger group which has been described in other publications (**18**, **19**), the *Ss* were all members of the high eighth and low ninth grades of a public junior high school in Oakland, California. For experimental purposes they were divided into two closely equivalent groups, with age characteristics as shown in Table I. The mean IQ (Kuhlmann-Anderson group test) was at this time approximately 102, with no significant differences between boys and girls.

TABLE I

Average Age of the Subjects ± $S.D._{Mean}$

Group	I		II	
	Boys	Girls	Boys	Girls
N	30	29	27	28
C.A.	14.1 ± .58	14.0 ± .55	14.0 ± .65	14.2 ± .73

Apparatus

The punchboard, in various forms one of the standard instruments for measuring eye-hand coördination and motor steadiness, has not yet been extensively utilized in a learning problem, so that the name does not appear in the index of the Psychological Abstracts. A multiple-hole punchboard can, however, be successfully applied to trial-and-error learning, as shown in a number of previous studies (**10**, **29**, **37**). As compared with the slot maze, it has the advantage of a simpler basic construction, and is also more readily adaptable to various experimental requirements. Since each choice point requires a discrete motor reaction, the learning process is relatively free from such performance factors as 'goal impetus' and 'centrifugal swing,' which may complicate maze learning. Within a normal range of *Ss*, learning is also unaffected by motor

coördination and precision, for although the punchboard was originally used as a motor test these factors can be eliminated as differentials by increasing the size of the holes in relation to the diameter of the stylus. In studying the rôle of informative signals, the punchboard has the advantage that right or wrong responses can be recognized only by these signals, whereas in the stylus maze the situation is complicated by other, kinesthetic, factors, that give direct indication of blind alleys and open pathways.

The apparatus used in this experiment was planned with the object of providing signals that would automatically register 'right' and 'wrong' responses, and that would also discriminate between different types of wrong responses.

Fig. 1 shows the externals of the instrument, as presented to the subject. The slanting

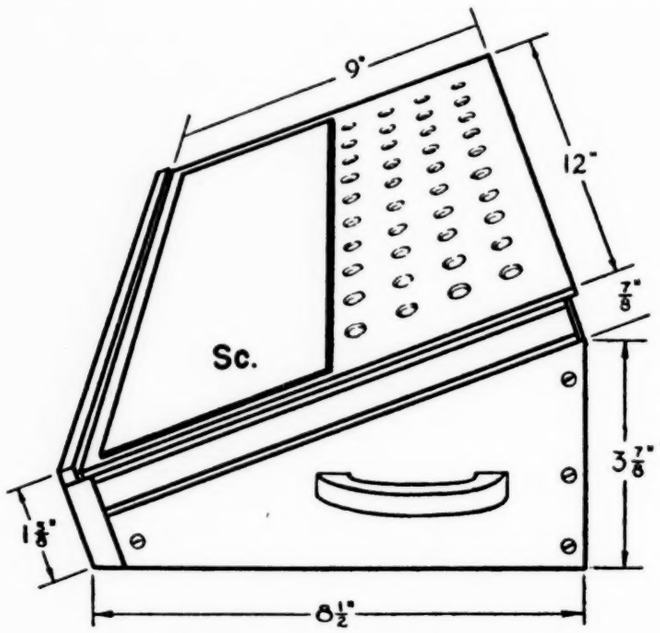

FIG. 1. Side view of punchboard.

top surface consists of a bakelite plate perforated with 10 columns of holes, four holes in each column. The holes are 3/8" in diameter, 1" between their centers. Beginning with the first column, or 'entrance,' the *S*'s task is to thrust a stylus through the holes until he discovers one which gives a signal previously specified as denoting the 'correct' hole. The other holes either give no signal, or one of several signals specified as indicating an error. When the correct hole is located, the *S* proceeds to the next column of four holes, and so on until he reaches the last or 'exit' column. This constitutes one trial. The procedure is repeated until two successive trials have been run without error, or (in a few exceptionally slow cases) until 15 trials have been completed. For each trial, a graphic record of responses and of time is obtained by means of a Bristol Recorder with five writing pens (one for each of four classes of error, and one for the correct response).

The differential signals are controlled by means of a contact device shown in Fig. 2. This is a removable bakelite plate which is inserted, like a drawer, in a slot at the side of the punchboard. Attached to the bakelite are rows of brass strips (Fig. 2) terminating in binding posts, with connections leading to a battery of stimulus producers. When inserted in the punchboard, the brass strips of pattern A lie directly below the rows of holes shown in Fig. 1. Thus, when a stylus is thrust through any hole in the top of the punchboard, it will complete an electric circuit by coming in contact with one of the brass strips, *except* at those points at which the brass is itself perforated, as indicated by the pattern of shaded circles in Fig. 2. Lying below the surface plate, these additional perforations are invisible to the *S*. Where they occur, the stylus will fall through to a bottom plate, and complete a circuit with the 'correct' stimulus. The *S*'s problem is to learn, as quickly as he can, to punch the pattern represented by these circles, and to avoid punching any of the other holes which will lead to 'error' signals. With this equipment, an experimental comparison can be made of learning curves based on different types of 'corect' signals, and the elimination of errors can also be examined in relation to the nature of the error indicators.

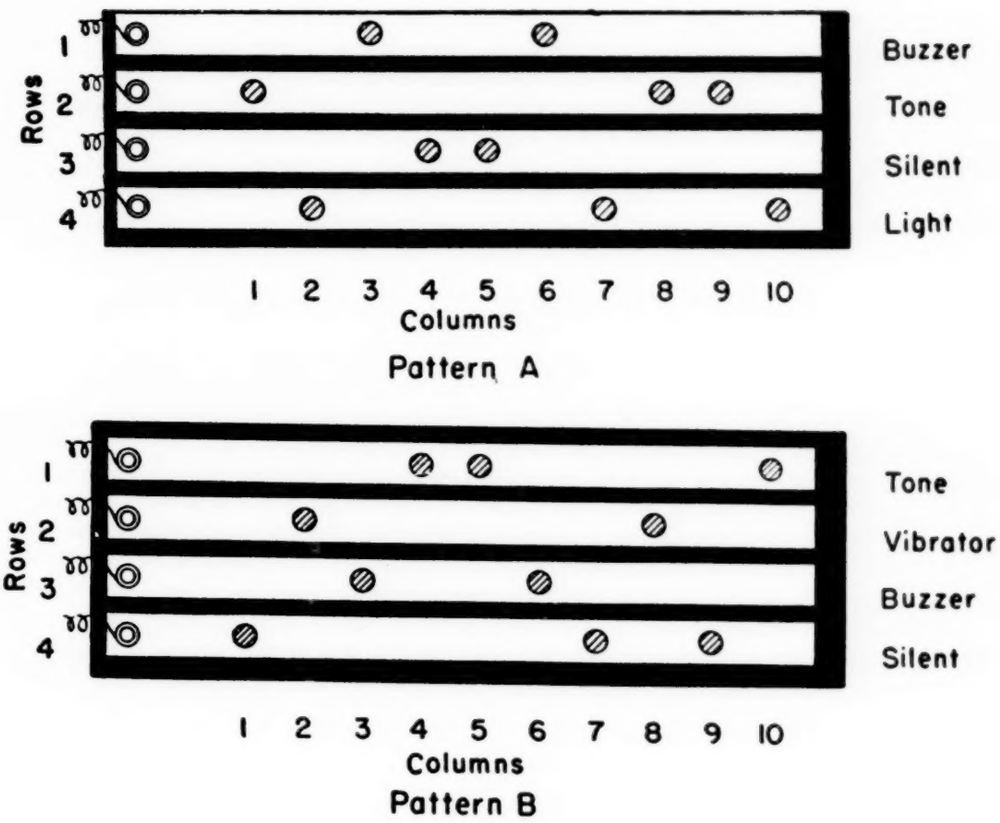

FIG. 2. Patterns A and B

EXPERIMENTAL PROCEDURE

Two patterns were used, as shown in Fig. 2. The second pattern was presented by shifting the screen (Sc., Fig. 1) from the lower to the upper half of the punchboard face plate. In balanced halves of the experiment, two stimuli were alternatively used to indicate the 'correct' response: a system of lights, and a 'vibrator' device in the handle of the stylus held by the *S*.[2] While it was not essential that the two patterns be of equal difficulty, they were set up so as to be at least closely similar in this respect.

The light was selected as a presumably pleasant stimulus, the vibrator as a disagreeable one. A preference contrast was confirmed by the *S*s who were asked to rate each stimulus on a 5-point scale, 1 signifying 'like it very much;' 2, 'like;' 3, 'indifferent;' 4, 'dislike;' and 5, 'dislike very much.' These ratings, made at the end of the experiment, and summarized in Table II, probably

TABLE II

SUBJECTS' RATINGS OF 'CORRECT' STIMULI

	Boys	Girls
Light	1.6	1.5
Vibrator	3.5	3.6

indicate a smaller difference than would be found in earlier stages of the experiment, since negative adaptation may be assumed to shift both stimuli toward the indifference point.

[2] The lights consisted of small red and green luminous points in a miniature Christmas tree, standing behind the apparatus and about eight in. in height. The vibrator consisted of a solenoid six in. long and one-half in. in diameter. The steel core of the solenoid, three-sixteenths in. in diameter, projected one in. beyond the coil and served as a stylus to penetrate the holes of the punchboard. When a circuit was established through a transformer, the stylus, and the *S*'s hand, were agitated at the rate of 60 vibrations per sec.

The scheme of presentation is shown in Table III.

TABLE III

	Pattern A (Order 1)	Pattern B (Order 2)
Group I	Correct cue: Light	Correct cue: Vibrator
Group II	Correct cue: Vibrator	Correct cue: Light

Results

Table IV shows the mean error scores for each group in each procedure.[3]

TABLE IV

Mean Errors ± S.D.Mean

		Group I Light Cue	Group II Vibrator Cue
Order 1 (Pattern A)	Boys	76.63 ± 4.77	76.89 ± 5.63
	Girls	82.38 ± 5.84	94.61 ± 8.18
		Vibrator Cue	Light Cue
Order 2 (Pattern B)	Boys	58.80 ± 3.54	59.33 ± 6.44
	Girls	57.07 ± 4.30	64.29 ± 7.67

Holding the order constant, it can be seen that in the case of boys the error scores are substantially identical, regardless of whether the guiding cue is the light or the vibrator. In the case of girls, more errors occur on the vibrator in order 1 and on the light in order 2. Since these small differences are consistently in favor of the girls in Group I, but in opposite directions so far as the stimuli are concerned, they are probably to be explained in terms of a sampling difference for the two groups of girls, rather than as due to any differential effectiveness of the stimuli involved.

Table V gives similar indications on the basis of time scores, with

TABLE V

Mean Time Scores in Seconds (± S.D.Mean)

		Group I Light Cue	Group II Vibrator Cue
Order 1 (Pattern A)	Boys	321.57 ± 14.37	344.22 ± 19.10
	Girls	357.34 ± 11.05	381.68 ± 27.16
		Vibrator Cue	Light Cue
Order 2 (Pattern B)	Boys	262.42 ± 14.27	248.59 ± 12.75
	Girls	263.66 ± 14.35	270.25 ± 14.20

no consistent nor reliable differences between the light cue and the vibrator cue, when these are compared (horizontally) with the order of presentation controlled. When vertical comparisons are made, in Tables IV and V, it can be seen that reliable differences occur

[3] Split-half coefficients, based on errors for odd and even trials and corrected for length of test, ranged from .89 to .95 in different parts of the experiment. Correlations between errors and time ranged from .58 to .79.

between error scores, and also between time scores, in the first and second orders of presentation. Assuming that the two patterns are similar in difficulty, we may conclude that there is a general practice effect in this experiment, manifested in a transfer from one pattern to another; the amount of transfer, however, is independent of the change in cues, for the gain is as great in shifting from a pleasant to an unpleasant cue, as it is when the change of cues is in the opposite direction.

Since, as previously noted, the difference in affective or preference values of the two stimuli may be greater in the earlier stages of the experiment, it may be desirable to examine the actual learning curves obtained under each of the two experimental conditions. These are shown in Fig. 3. Since no significant differences can be demonstrated at any point, it would appear that in this experiment the

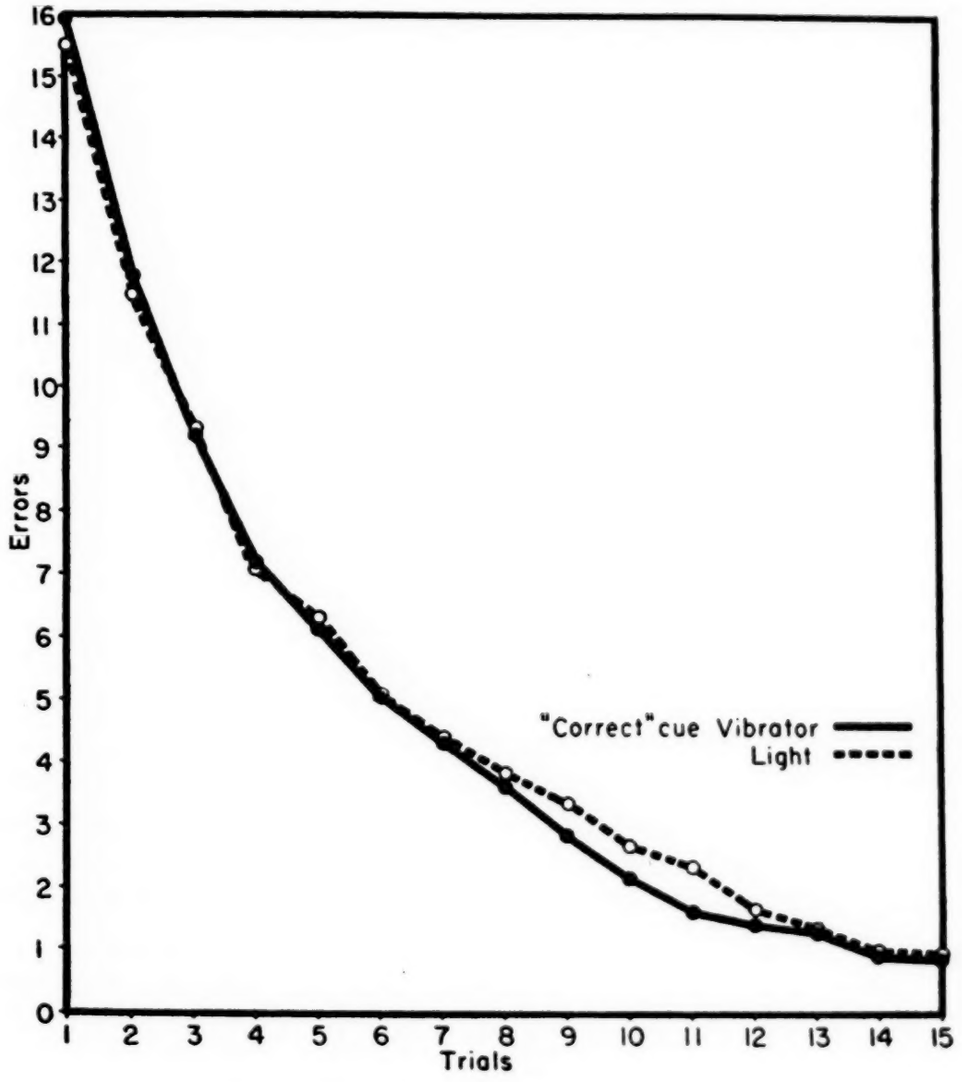

FIG. 3. Punchboard learning curve (errors).

goal of learning a pattern is of predominant importance, and that, within the range of stimulus values which we have employed, the intrinsic pleasantness or unpleasantness of the cues is of little or no significance in the learning process. In one of the earlier statements of the law of effect, Thorndike proposed that "The greater the satisfaction . . . the greater the strengthening of the bond" (**33**, p. 244). When, however, we consider the satisfaction attendant upon an act, in terms of greater or less, it may be necessary to restrict our statement of the law of effect to those affective or motor-affective conditions which are directly relevant to the responses to be learned; if an *S* is motivated to perform efficiently, the satisfaction of a 'right' response is the primary and possibly the only affective factor which is directly linked with the learning of this response.[4] The incidental affective values of the cue stimuli do not appear to be additive to this primary affective factor. In experiments in which more intense stimuli have been used (e.g., electric shock), the positive or negative consequences sometimes noted in the learning process perhaps occur chiefly through 'alerting' or disrupting performance, rather than through mechanisms that can be hypothesized for the law of effect.

The Elimination of Errors

A comparison of the differential effect of stimuli can also be made with reference to the elimination of errors. In an earlier report from this laboratory, dealing with results with a stylus maze, it was noted: "To a casual observer the . . . experiment offers striking contrasts in procedure. The physical factors (external to the maze) differ widely in affective intensity. Unquestionably they operate in determining how much a person enjoys the learning task, and in determining his readiness to undertake similar tasks, but they have no effect upon the learning process itself" (**20**, p. 471). In this previous experiment errors were signalled by different methods in successive experimental procedures. In the present study, the construction of the punchboard makes it possible to use differential signals *within* a given procedure, and to examine the rate of elimination of errors according to the signals used to define them. The error signals used included the light (when not used as a 'correct' signal); the vibrator (when not used as a 'correct' signal); a buzzer; a radio signal tone produced through an audio-oscillator and also a 'silent'

[4] In Thorndike's more recent statement "A satisfying after-effect which *belongs* to a connection can be relied on to strengthen the connection" (**34**, p. 311, italics ours). The concept of 'diminishing returns' (**36**) may also deserve consideration here. The negative results reported above may come either from the irrelevance ('not belonging') of the experimental stimulus changes, or from the fact that at this motivational level any added increments are too small to register.

signal. Table VI indicates the experimental procedure, and the type of stimulus given when a response in any row is an error.

TABLE VI

ERROR CONNECTIONS, BY ROWS

		Group I	Group II
Order 1 (Pattern A)	Row 1	Buzzer	Buzzer
	Row 2	Tone	Tone
	Row 3	Silence	Silence
	Row 4	Vibrator	Light
	Correct Signal	Light	Vibrator
Order 2 (Pattern B)	Row 1	Tone	Tone
	Row 2	Light	Vibrator
	Row 3	Buzzer	Buzzer
	Row 4	Silence	Silence
	Correct Signal	Vibrator	Light

Table VII shows the affective values of the error signals, as rated by the *S*s on a five-point scale from 'like very much' (*1*) to 'dislike very much' (*5*).

A fair agreement can be noted in the ratings by boys and girls, except that the girls on the average were indifferent to the radio

TABLE VII

SUBJECTS' RATINGS OF ERROR SIGNALS

Stimulus	Average Rating	
	Boys	Girls
Light	1.6	1.5
Tone	2.1	3.0
Silence	2.4	2.2
Buzzer	2.7	3.5
Vibrator	3.5	3.6

signal tone and disliked the buzzer, while the boys tended to like all of the signals except the vibrator.

Although the performance of boys was slightly superior to that of girls, the difference was not reliable, and data on the elimination of errors will be presented for the two sexes combined. Table VIII shows the total errors, and the standard deviation of the mean, according to group, pattern, and the specific stimulus factor.

When horizontal comparisons are made in Table VIII, it will be noted that for a specific stimulus within a given pattern the mean error scores are closely similar in Groups I and II. Thus, in Pattern A the mean error score on holes signalled by the buzzer is 15.85 for Group I and 16.41 for Group II. A considerable difference can be noted between different stimuli on a given pattern. Thus, in Pattern *A*, Group I, the critical ratio is 7.6 between mean errors on the

TABLE VIII

MEAN ERRORS ON DIFFERENT STIMULUS CUES (± S.D.$_{Mean}$)

		Group I Order 1 Pattern A	Group II Order 1 Pattern A
Row 1	Buzzer	15.85 ± .93	Buzzer 16.41 ± 1.13
Row 2	Tone	26.83 ± 1.49	Tone 29.11 ± 1.69
Row 3	Silence	23.85 ± 1.23	Silence 25.93 ± 1.56
Row 4	Vibrator	12.97 ± .82	Light 14.16 ± .98
		Group I Order 2 Pattern B	**Group II Order 2 Pattern B**
Row 1	Tone	12.36 ± .75	Tone 14.09 ± 1.09
Row 2	Light	19.41 ± 1.07	Vibrator 21.65 ± 1.51
Row 3	Buzzer	15.22 ± .76	Buzzer 16.24 ± 1.11
Row 4	Silence	10.25 ± .74	Silence 9.91 ± .75

buzzer and on the tone, and 11.7 between mean errors on 'silence' and on the vibrator. In each case, the number of errors is reliably smaller for the less pleasant, or more 'punishing' stimulus. A further inspection of the table, however, and a comparison of the same stimuli in different patterns, suggests that the differentiating factor is not the stimulus, but the row of the punchboard on which the stimulus falls. Regardless of the stimulus, errors on Rows 1 and 4 are relatively easy to eliminate. These are, respectively, on the upper and lower margins of the punchboard; this is particularly true of Row 4, which is nearest to the *S*. Errors on the middle rows, particularly on Row 2, are less easy to eliminate. The reason for this is not difficult to conceive; in fixating characteristics of a hole in, for example, Row 1, the *S* may easily confuse this with the adjacent hole in Row 2. But a hole in Row 2 may be confused with adjacent holes in either Row 1 or Row 3. In terms of the percentage of total errors, 20.9 percent fall on Row 1, 34.2 percent on Row 2, 28.3 percent on Row 3, and 18.7 percent on Row 4.

The characteristics of specific stimuli, in relation to learning, can be examined more readily if the mean errors are translated into terms of the percentage of total errors for that particular part of the experiment; each stimulus is then represented once in an inner or easier row, and once in an outer or more difficult row, and averaged so as to reduce the influence of spatial position. The results are shown in Table IX.

TABLE IX

AVERAGE PERCENTAGE OF TOTAL ERRORS

Light	(Rows 2 and 4)	24.0
Tone	(Rows 2 and 1)	28.0
Silence	(Rows 3 and 4)	23.6
Buzzer	(Rows 3 and 1)	23.0
Vibrator	(Rows 2 and 4)	25.6

The proportions of total errors, for the various stimuli, do not differ significantly from the value of 25 percent which would be expected on a chance basis, nor do they differ significantly from each other. It is clear that no tendency exists for the more unpleasant stimuli, or for the more intense stimuli, to produce a more rapid elimination of the errors which they serve to indicate. These errors *are* eliminated, but the mechanism appears to be unrelated to the physical characteristics and affective qualities of the stimuli through which the errors are recognized.

One uncontrolled feature of this experiment is that for each group the error stimulus in Row 2, Pattern B, has previously served as a correct signal in Pattern A. This could conceivably result in increased errors in Pattern B for the stimuli to which positive reactions have previously been given. Because of incomplete balancing with regard to rows and columns this factor cannot be examined directly. However, it can be noted that for Groups I and II combined, the percentage of total errors in Row 2 is substantially the same in the first and in the second orders (33.8 percent and 33.4 percent respectively). This suggests that under changed instructions as to the correct stimuli, there is no transfer of response to specific stimuli from one experiment to the next.

The results from error analysis are consonant with those of Muenzinger, who presents an able discussion of differences between rats and humans in their response to emphasizing factors. Although Muenzinger (**24**) found that electric shock apparently provides an additional motivation in rats, this was not true of human subjects (**25**), who were performing in response to quite different drives and were already so well motivated that the addition of shock had no effect.

The argument can perhaps be offered that since there is no differential stamping-out of errors in relation to the cue stimuli, the process of error elimination should not be regarded as 'piecemeal,' but as representing some more complex form of seizing the pattern 'as a whole.' This, however, is not a necessary conclusion, for the right choices may be confirmed and the wrong choices eliminated on a piecemeal basis solely through the reactions attending the knowledge of 'right' and 'wrong.' Although the present experiment was not set up to examine this particular problem, information of at least incidental value may be obtained from an analysis of errors in relation to position to series. In learning correct choices a difference is of course to be expected between the columns as well as between the rows, in that a peripheral position is more favorable than one nearer the middle of the series. This is illustrated in Fig. 4, which presents the average errors per column (per choice point) for the two

patterns combined and for each of four sub-groups segregated on the basis of total error score. Group A, in this figure, consists of the best learners, with total errors one S.D. or more below the group mean; B, C and D represent zones of the distribution with progressively larger error scores (0 to − 1 S.D.; 0 to + 1 S.D.; + 1 S.D. or more). Since errors at a given choice point are influenced not merely by position in the pattern but also by purely local factors,[5]

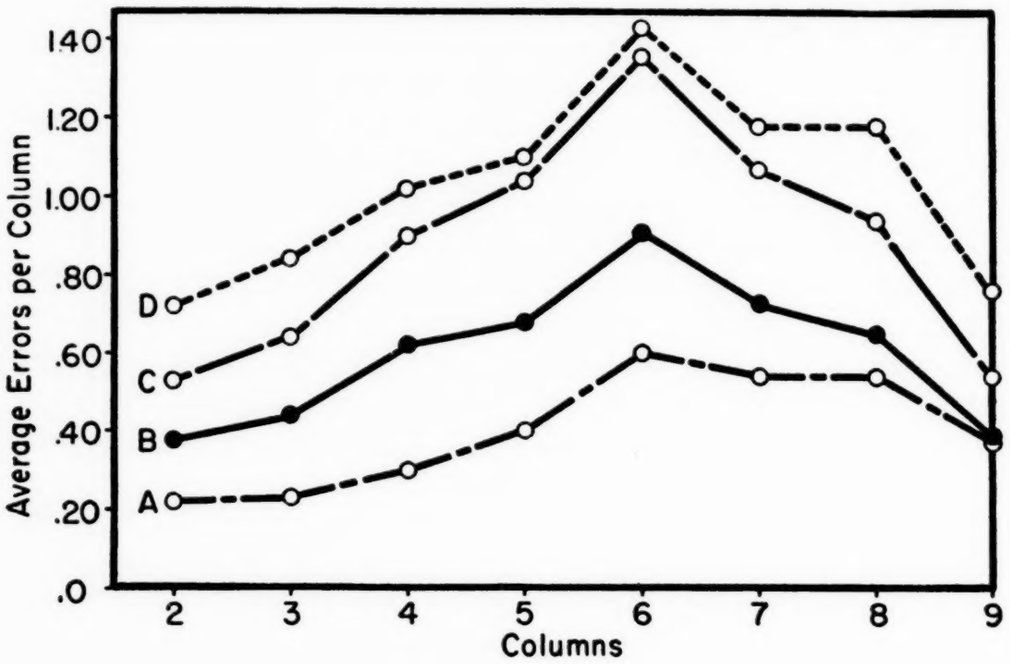

FIG. 4. Punchboard errors, by columns, in four ability groups.

these have been eliminated to some extent in Fig. 4, by computing a three-point running average.

As in other analyses of intra-serial phenomena (**23**), Fig. 4 shows an increase in difficulty, continuing to a point slightly beyond the middle of the series, and then diminishing toward the end. The error curves for the four subgroups are, moreover, nearly parallel, showing that error elimination follows a similar course at different ability levels. So far as this evidence goes we may conclude that efficient and inefficient performers on the punchboard are distinguished quantitatively rather than qualitatively. The good learner is apparently not characterized by sudden integrations of the total pattern, embracing all choice points, but (like average and poor learners) by the analytic, piecemeal elimination of errors. This

[5] Thus, it is more difficult to eliminate errors if the correct choice occurs twice in succession on the same row. It also appears more difficult if the correct choice shifts diagonally downward than if it shifts diagonally upward.

process is relatively rapid in the good learner, but it involves the same difficulty pattern as in the poor learner.

Summary

1. An apparatus is described for the study of trial-and-error learning, consisting of a punchboard with graphic recording of responses and time. Two patterns are provided, with differential signals for right responses and for different classes of errors.

2. A learning experiment was conducted with a group of 114 junior high school children, equally divided between boys and girls, and subdivided into two experimental groups.

3. Learning was compared under conditions in which the correct responses were signalled by stimuli of differing affective quality or preference value. The degree to which the cue stimuli were 'liked' or 'disliked' was found to bear no relation to average time, average errors, or to the course of the learning curve. No gain can be expected from 'sugar coating' a correct response, by providing a supplementary pleasant experience to accompany the response.

4. Learning was compared under four experimental variations, with five different signalling conditions for different classes of errors. The rate of elimination of errors was found to be unrelated to the nature of the stimulus used to signal errors. In the avoidance of specific choices, no differences can be expected from the differential accenting or punishing of these choices; the mere knowledge that a response is wrong is, on the average, as effective as this knowledge plus an emphasizing stimulus.

5. The elimination of errors was markedly influenced by position in the pattern.

6. In a comparison of good and poor learners, the elimination of errors was found to follow a similar course. This would be expected if learners at different ability levels are similarly dependent upon piecemeal trial-and-errors in the learning process.

(Manuscript received May 24, 1944)

References

1. Bernard, J., & Gilbert, R. W. The specificity of the effect of shock for error in maze learning with human subjects. *J. exp. Psychol.*, 1941, **28**, 178–186.
2. Bunch, M. E. Certain effects of electric shock in learning a stylus maze. *J. comp. Psychol.*, 1935, **20**, 211–242.
3. Cole, L. W. The relation of strength of stimulus to rate of learning in the chick. *J. Anim. Behav.*, 1911, **1**, 111–124.
4. Courts, F. A., & Waggoner, D. The effect of 'something happening' after a response. *J. exp. Psychol.*, 1938, **22**, 383–387.
5. Crafts, L. W., & Gilbert, R. W. The effect of punishment during learning upon retention. *J. exp. Psychol.*, 1934, **17**, 73–84.

6. Dodson, J. D. Relative value of reward and punishment in habit formation. *Psychobiol.*, 1917, **1**, 231–276.
7. Dorcus, R. M., & Gray, W. L. The effectiveness of food and electrical shock in learning and retention by rats when applied at critical points in the maze. *J. comp. Psychol.*, 1932, **14**, 191–218.
8. Drew, G. C. The function of punishment in learning. *J. genet. Psychol.*, 1938, **52**, 257–267.
9. Eisenson, J. Confirmation and information in rewards and punishments. *Arch. Psychol., N. Y.*, 1935, **181**. Pp. 37.
10. Farber, I. Spread of effect of reward and punishment in a multiple choice situation. *Proc. Ia. Acad. Sci.*, 1941, **48**, 313–317.
11. Forlano, G. School learning with various methods of practice and rewards. *Teach. Coll. Contr. Educ.*, 1936, No. 688. Pp. 114.
12. Gilbert, R. W. The effect of non-informative shock upon maze learning and upon retention with human subjects. *J. exp. Psychol.*, 1936, **19**, 456–466.
13. Gilbert, R. W., & Crafts, L. W. The effect of signal for error upon maze learning and retention. *J. exp. Psychol.*, 1935, **18**, 121–132.
14. Gurnee, H. The effect of electric shock for right responses on maze learning in human subjects. *J. exp. Psychol.*, 1938, **22**, 354–364.
15. Hoge, M. A., & Stocking, R. J. A note on the relative value of punishment and reward as motives. *J. Anim. Behav.*, 1912, **2**, 43–50.
16. Holodnak, H. B. The effect of positive and negative guidance upon maze learning in children. *J. educ. Psychol.*, 1943, **34**, 341–354.
17. Hulin, W. S., & Katz, D. A comparison of emphasis upon right and wrong responses in maze learning. *J. exp. Psychol.*, 1935, **18**, 638–642.
18. Jones, H. E. Procedures of the adolescent growth study. *J. consult. Psychol.*, 1939, **3**, 177–180.
19. Jones, H. E. *Development in adolescence.* New York: D. Appleton-Century Co., 1943. Pp. xvii + 166.
20. Jones, H. E., & Yoshioka, J. A. Differential errors in children's learning on a stylus maze. *J. comp. Psychol.*, 1938, **25**, 463–480.
21. Kellogg, W. N. Electric shock as a motivating stimulus in conditioning experiments. *J. gen. Psychol.*, 1941, **25**, 85–96.
22. Lorge, I. The efficacy of intensified reward and of intensified punishment. *J. exp. Psychol.*, 1933, **16**, 177–208.
23. McGeoch, J. A. *The psychology of human learning, an introduction.* New York: Longmans, Green & Co., 1942. Pp. xvii + 633.
24. Muenzinger, K. F. Motivation in learning. I. Electric shock for correct response in the visual discrimination habit. *J. comp. Psychol.*, 1934, **17**, 267–277.
25. Muenzinger, K. F. Motivation in learning. II. The function of electric shock for right and wrong responses in human subjects. *J. exp. Psychol.*, 1934, **17**, 439–448.
26. Porter, E. H., Jr. An investigation of the alleged function of emphasis in a simple discrimination problem. *J. exp. Psychol.*, 1941, **28**, 77–84.
27. Porter, E. H., Jr., & Hall, C. S. A further investigation of the role of emphasis in learning. *J. exp. Psychol.*, 1938, **22**, 377–383.
28. Rock, R. T., Jr. The influence upon learning of the quantitative variation of after-effects. *Teach. Coll. Contr. Educ.*, 1935, No. 650. Pp. xii + 78.
29. Silleck, S. B., Jr., & Lapha, W. The relative effectiveness of emphasis upon right and wrong responses in human maze learning. *J. exp. Psychol.*, 1937, **20**, 195–201.
30. Stephens, J. M. The influence of different stimuli upon preceding bonds. *Teach. Coll. Contr. Educ.*, 1931, No. 493. Pp. vii + 83.
31. Stephens, J. M. The influence of symbolic punishment and reward upon strong and upon weak associations. *J. gen. Psychol.*, 1941, **25**, 177–185.
32. Stephens, J. M., & Baer, J. A. The influence of punishment on learning when the opportunity for inner repetition is reduced. *J. genet. Psychol.*, 1937, **51**, 209–217.
33. Thorndike, E. L. *Animal intelligence.* N. Y.: Macmillan, 1911. Pp. viii + 297.
34. Thorndike, E. L. *The fundamentals of learning.* New York: Bureau of Publications, Teachers College, 1932. Pp. xvii + 638.
35. Thorndike, E. L. Reward and punishment in animal learning. *Comp. Psychol. Monog.*, 1932, **8**, No. 4. Pp. 65.

36. THORNDIKE, E. L., & FORLANO, G. The influence of increase and decrease of the amount of reward upon the rate of learning. *J. educ. Psychol.*, 1933, **24**, 401–411.
37. TOLMAN, E. C., HALL, C. S., & BRETNALL, E. P. A disproof of the law of effect and a substitution of the law of emphasis, motivation, and disruption. *J. exp. Psychol.*, 1932, **15**, 601–614.
38. TOMKINS, S. S. An analysis of the use of electric shock with human subjects. *J. Psychol.*, 1943, **15**, 285–297.
39. TRAVIS, R. C., & ANDERSON, H. C. The effect of electric shock on learning in eye-hand coördination. *J. exp. Psychol.*, 1938, **23**, 101–107.
40. TUCKMAN, J. The influence of varying amounts of punishment on mental connections. *Teach. Coll. Contr. Educ.*, 1933, No. 590. Pp. v + 45.
41. WARDEN, C. J., & AYLESWORTH, M. The relative value of reward and punishment in the formation of a visual discrimination habit in the white rat. *J. comp. Psychol.*, 1927, **7**, 117–127.

EXPERIMENTAL STUDIES IN AFFECTIVE PROCESSES: II. ON THE QUANTIFICATION AND EVALUATION OF 'MEASURED' CHANGES IN SKIN RESISTANCE

BY ERNEST A. HAGGARD

Harvard University

If one is going to make measurements, he should keep at least two things in mind: he should know, first, the characteristics of his 'measuring stick,' and second, that any generalizations he may choose to make on the basis of his 'data' will be meaningful in direct proportion to the precision of his system of quantification. Or, stated differently, one cannot, by the mere assignment of numerical values to extensive qualities in the empirical world *ipso facto* provide the 'measured' qualities with all of the characteristics of the numerical system or model that he employs. Hence it is often necessary to consider certain methodological problems before one can comprehend the significance of specific experimental findings. In this connection, the purpose of the present paper is (*a*) to examine some of the problems involved in the quantification and evaluation of recorded changes in skin resistance, and (*b*) to propose a method which appears to satisfy certain basic tenets of mensuration.

As in most types of research, work in this field is subject to certain conditions which may distort the data if not properly accounted for. These conditions may be thought of as possible sources of error,[1] and include the nature of the apparatus, individual differences among the *S*s, and the evaluation and interpretation of the obtained data. These sources of error have already received more or less attention in the literature (e.g., **5, 6**), and, except for a brief mention of the importance of variations in apparatus and *S*s, only the general problems involved in the evaluation of the raw data will be treated in this paper.

Variations due to differences in apparatus.—In the measurement of changes in skin resistance, the nature of the results is dependent on such factors as the basic type of apparatus used (e.g., Wheatstone bridge or vacuum tube type), its sensitivity and characteristic period, and the degree of non-linearity in its measurements (i.e., whether there is a systematic distortion in the magnitude of the deflections corresponding to ohmic changes along the total resistance range), the size and type of electrodes, the bodily areas from which the records were taken (e.g., the palm or back of the hand), the temperature of this area, whether the external current passing through the subject is A.C. or D.C., the size of this current, and whether it remains constant or varies with the *S*'s change in resistance. Such a list of items to be considered may be somewhat disconcerting at first glance, but the investigator must not disregard them simply because they complicate his task.

[1] 'Error' is used here in a statistical as well as in an experimental sense—that is, as variations in the array of scores which are due to uncontrolled or unknown factors. These variations go to make up the denominator in a significance ratio. Unfortunately, however, many of the sources of error mentioned in this paper act in a constant rather than in a random manner and so should be controlled.

Variations due to differences among subjects.—Several studies have reported individual differences among *S*s which one may or may not need to consider. These include the *S*'s state of health or nutrition, his over-all emotional adjustment or maladjustment, his feelings of stress or anxiety in the experimental situation, his mental set, and the extent of his general muscular tonus or relaxation. Variations in any or all of these conditions may help to determine the nature of the galvanic records.

Evaluation of the raw data.—Assuming that variations due to differences in the characteristics of the apparatus and *S*s are properly accounted for, we come to the problem of how the raw data should be quantified. Some degree of quantification is of course necessary, unless one is satisfied to make only the most casual statements. For example, it is not very satisfactory just to say, "I put a subject in the apparatus, presented the stimulus, and the needle moved 'this much.'" One must know a good deal about the significance of 'this much.' Furthermore, the raw data or 'experimental findings' cannot necessarily be taken at face value—even though they may appear in objective or physical terms.

Several methods of representing changes in skin resistance have been proposed in the literature. Included are such measures as: (*a*) whether or not a deflection of the indicator was observed (10), (*b*) the absolute change in ohms resistance (10), (*c*) the distance that the indicator moved—in terms of millimeters or inches (10), (*d*) the percentage that a given deflection, or GSR, is of the *S*'s total range of deflections (10), (*e*) the percentage that a given deflection is of the *S*'s general resistance level (1, 10), (*f*) the deflection translated into standard measures (*z*-scores) on the basis of the *S*'s total distribution of deflections (11), (*g*) the deflection translated into a change in conductance or micromho[2] units (2, 4, 7, 10), and finally, (*h*) the deflection translated into a change based on the logarithm of the conductance, or log. conductance units (5, 8). Which of these scales is most suitable for the treatment of the raw data? Or does it make much difference which one is used? And, if it does make a difference, on what grounds should the selection be made?

But before examining certain scales which have been proposed as methods of quantifying skin resistance data, it seems clear that any method which is really satisfactory should possess certain demonstrable properties. More specifically, the method used should enable one:

1. To make the necessary mathematical transformations with a maximal degree of simplicity and a minimal number of computational errors.
2. To speak of individual differences among *S*s under comparable experimental conditions as well as differences among groups of *S*s receiving variations in experimental treatments.
3. To speak of the general level of background excitation (general level) independently of the particular reaction (GSR) which, as a result of a specific stimulus or set of stimuli, is superimposed on the preëxisting level of excitation.
4. To use such statistical techniques as product-moment correlation, *t*-tests, analysis of variance, etc. In using these methods of analysis, one makes the implicit assumption that his scale of measurement possesses units of equal size over its total range.

There has obviously been a good deal of disagreement over which of the scales listed above is 'the best,' as evidenced by their number and variety. Probably the most straightforward measure is a record of the change in the absolute number of ohms (or the magnitude of the needle's deflection, which amounts to the same thing). Several writers have considered this procedure adequate.[3] But it has been apparent for some time that the size of the GSR in terms of ohms change is intimately related to the subject's general level of resistance. Wenger and Irwin (11), for example, found a correlation of .90 between the magnitude of the GSRs (ohms) and the general level (ohms), thus clearly showing that GSRs from different resistance levels cannot be compared directly. They got around this difficulty by translating their data into standard measures

[2] Since conductance here is defined as the reciprocal of resistance ($C = 1/R \times 10^6$), the term which designates conductance (mho) is obtained by spelling the measure of resistance (ohm) backwards.

[3] Even though Hunt and Hunt (10) concluded that the use of the absolute number of ohms change seems preferable to other methods of measurement on the basis of the intercorrelations among these measures which approach unity (*cf.* proposals *a*, *b*, *c*, *d*, *e* and *g* above), it is clear from Fig. 1 that these high positive correlations could not have been obtained if their *S*s had displayed a substantial change in their general level of resistance.

(z-scores).[4] Other attacks on this problem were made by Darrow. On the basis of data which demonstrated that the amount of perspiration tends to vary as the reciprocal of the resistance, or as conductance (3, 4), he suggested the use of conductance or micromho units as an indication of the concomitant general physiological and psychological activity accompanying the GSR (2, 4). Later, however, Darrow (5) altered this position on the grounds that the nature of psychophysical 'laws' renders it unlikely that measured changes in local physiological activity bear a linear relation to the magnitude of psychological change, and instead suggested the use of a measure based on the logarithm of the concomitant physiological activity, or on log. conductance units.

The question which then arose was concerned with the nature of the empirical properties of the proposed scales. That is, it seemed important to know the characteristics of the functional relationship (if any existed) between the two measures of galvanic activity: namely, the general level of skin resistance (and the inferred level of background excitation) and the momentary decrease in the skin resistance or the GSR (and the inferred increment in background activity as a result of a particular stimulus). Three scales were selected for study in this connection: the ohms scale, because the basic or raw data appeared in this form, and the Darrow's conductance scale and the log. conductance scale, because of their foundation in certain physiological data and theoretical formulations.

Method

Basic study

Subjects.—48 young men of college age served as *S*s.

Apparatus.—The basic recording instrument was a G.E. photoelectric recording microammeter placed in series with a modified vacuum tube voltmeter (Junior Volt Ohmyst). The apparatus was so calibrated that both the general level of the *S*s' resistance, and the magnitude (in ohms) of all deflections could be determined immediately from the continuous ink records.[5]

Stimuli.—A list composed of the following 32 verbal stimuli were used: *city*, *pasture*, *taxi*, *hay*, *country*, *farm*, *easy*, *horse*, *library*, *barn*, *stores*, *soft*, *pavement*, *cafe*, *meadow*, *weak*, *carrot*, *traffic*, *subway*, *tender*, *sidewalk*, *corn*, *quiet*, *boulevard*, *harvest*, *office*, *poultry*, *theater*, *cow*, *kind*, *streetcar*, and *plow*. It will be seen that these words varied in several respects: 13 were rural terms, 13 were urban terms, and 6 were ambiguous, but may have been taken as terms of personal reference. This list of stimuli was considered satisfactory in so far as it covered a fairly wide range in the degree of familiarity, meaningfulness, etc., of the words.[6]

Experimental procedure.—After the *S* was comfortably seated in a chair, the zinc cup electrodes were attached to the palms of his hands.[7] These electrodes were then enclosed in small wooden boxes which were strapped to the *S*'s palms in order to minimize any changes in the position or pressure of the electrodes. The *S* was then told that the test would not start for about five min., during which time he might as well sit quietly and relax. This interval was used to let the *S* get used to the apparatus and general situation, and to allow time for polarization of the skin area adjacent to the electrodes.

---

[4] But by using this procedure, Wenger and Irwin were unable to make comparisons between subjects in terms of the relative size of their GSRs, since their standard measures were defined as the relative deviation from the mean of a specific distribution—i.e., the distribution of an individual *S*.

[5] It is worth noting in passing that the instrument was practically linear over the entire range—that is, when a resistance box was used for the initial calibrations, there was no observable distortion in the magnitude of the (recorded) deflections between 1000 and 30,000 ohms, although there was a distortion of approximately .023 at the 60,000 ohm level. An error in measurement this slight was considered negligible, and no correction was made for it. For a more complete description of this apparatus, see 8 and 9.

[6] Although a strong electric shock was administered to the *S*s during a subsequent session or a later part of the same session from which the responses were taken, none of the values comprising Fig. 1 were taken from *S*s who had received any electric shocks.

[7] Because all of the data reported in this paper are concerned with changes in the rate of palmar resistance, any generalizations are meant to apply to this type of data.

As was described in a previous paper (**8**), the *S* responded to each stimulus by: (*a*) a series of chain associations for an interval of 10–12 sec., and (*b*) by pressing a plunger with the index finger of the right hand at the time each association was made. A period of 20–40 sec. elapsed after the *S*'s last association to a particular word before the next stimulus word was presented.

Treatment of the data.—A total of 675 responses were recorded, and these comprise the data which are represented in Figs. 1 and 2. For each response, a measure was obtained of the *S*'s general level of resistance just before the stimulus was presented, the difference between this value and that of the resistance level obtaining a few seconds after the stimulus was presented, and the corresponding differences between the two measures from the conductance and log. conductance scales. More specifically, let us suppose that an *S*'s general level of resistance before the presentation of a given stimulus is 12,200 ohms, and that there is a decrease in this level of 800 ohms (or a drop of the level to 11,400 ohms) 'as a result of the stimulus.' The value of the ohms GSR (0.8) is obtained by subtracting the second from the first resistance level (i.e., 12.2 − 11.4 = 0.8 = ohms GSR).[8] Similarly, as will be seen in Table I, the corresponding conductance

TABLE I

(Sample Calculation of the GSR)

TABLE OF VALUES FROM WHICH ARE COMPUTED THE GENERAL LEVEL OF RESISTANCE AND THE GSRs FOR THE OHM, CONDUCTANCE, AND LOG. CONDUCTANCE SCALES

	1000 Ohm Values	Conductance Values	Log. Cond. Values
	—	—	—
(*a*) Reading taken at the time the stimulus was presented, which becomes the measure of the *Resistance Level*	12.3	—	—
	12.2	81.96	191.36
	12.1	—	—
	12.0	—	—
	11.9	—	—
	11.8	—	—
	11.7	—	—
	11.6	—	—
(*b*) Reading of maximal decrease in resistance occurring a few sec. after the stimulus was presented	11.5	—	—
	11.4	87.71	194.31
	11.3	—	—
	—	—	—
GSR: the absolute difference between readings *a* and *b*	0.8	5.85	2.95

and log. Conductance GSRs are obtained in the same manner. Thus, while the GSR from the ohms scale which represents this change in palmar skin resistance is 0.8 units, the GSR from the conductance scale is 5.85 units, and from the log. conductance scale is 2.95 units. This procedure was followed to obtain the resistance level and GSR values for each of the 675 reactions.

The scores for the general level of resistance (ohms) and the GSRs for these three scales were then punched into International Business Machine cards. These cards were later sorted (or arranged) according to resistance level (e.g., 12.2 in Table I), and the GSR values occurring at each resistance level were averaged for each of the three scales. This procedure was followed on the assumption that if a general 'law' or functional relationship existed between the resistance level and the size of the respective GSRs, it should appear when the individual differences in the reactional tendencies of the *S*s and the differences in the stimulus value, or significance of the stimuli, were randomized. For simplicity of presentation, the data were later grouped into 5000 ohm intervals, and are so depicted in Fig. 1. That is, the value of 1.5 for the ohms scale at the 25,000 ohm level was obtained by averaging all the GSR scores associated with a range of the general level falling between the 22,600 and the 27,500 ohms.

[8] It will be noted here that the value 12.2 (thousand ohms) corresponds to 12,200 ohms, as it does in Table I and Figs. 1 to 3. The translation of the resistance values to conductance and log. conductance values in Table I, however, are made from the original ohmic readings.

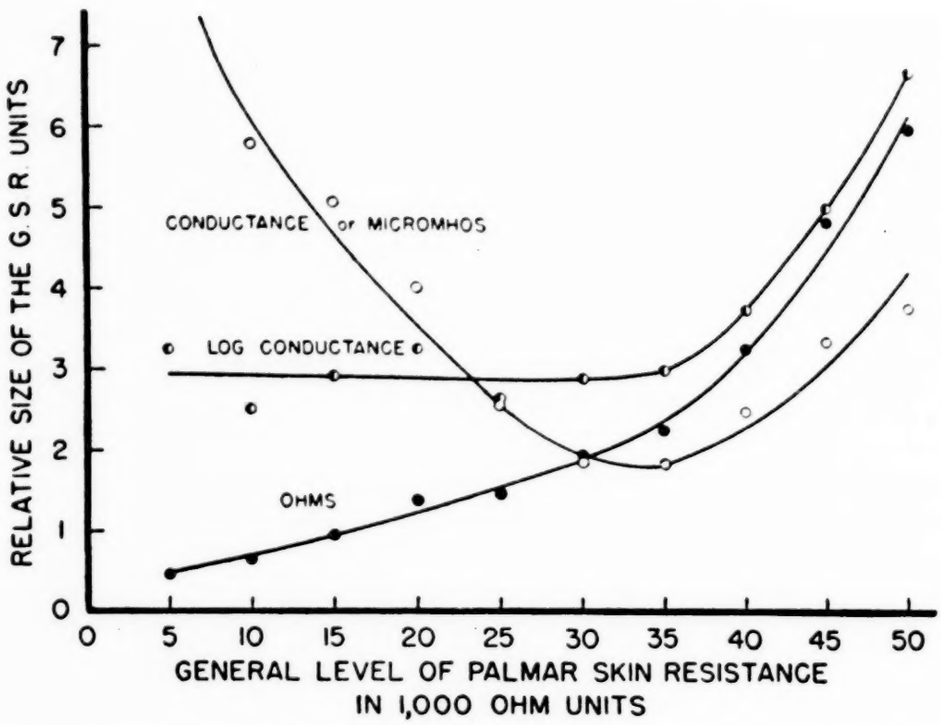

FIG. 1. A comparison of the relative size of the ohms, the conductance, and the log. conductance GSRs to 32 verbal stimuli in relation to the general level of resistance. (N = 675, with 48 *Ss*.)

Examination of the three scales in Fig. 1:

Ohms.—It is readily apparent that a functional relationship exists between the general level of skin resistance and the magnitude of the decrease in ohms resistance, or the size of the GSR in ohms units. Thus, under the conditions of the present study, we may say that—other things being equal—if an *S*'s general level of resistance is 15,000 ohms, his average GSR (in ohms units) will be about twice as great as it would be if his general level were 5000 ohms, but only about one-sixth as large as it would be if his general level were 50,000 ohms. This relationship, furthermore, is fairly stable considering the fact that these values are averages of the individual scores of 48 *Ss*, and are reactions to 32 different verbal stimuli. It should be pointed out, however, that one should not extrapolate beyond the values presented in Fig. 1, since it appears that at the upper resistance ranges (e.g., 40,000–50,000 ohms level) the curve starts to level out, or to fall along a sigmoid function.

Conductance.—The conductance scale, which is defined as the reciprocal of the ohms scale (times the constant, 10^6), obviously overcorrects for the shortcomings of the ohms scale in the lower resistance ranges by making the size of these intervals too large, but fails to provide sufficient correction in the higher resistance ranges.

It may be noted that the point for the conductance scale at the 5000 ohm level does not appear in Fig. 1. This is because the value was 11.7 conductance units, and hence did not fit on the graph.

Log. conductance.—Of the three scales represented in Fig. 1, the log. conductance scale is clearly the one to be preferred. That is, for the range between about 5000 and 35,000 ohms, the scale units are of approximately equal size, and hence meet the important qualifications listed above (see p. 47). On the basis of the data reported in this paper, it is an adequate translation for about three-fourths of the measures, or for those falling at or below the 35,000 ohm level.

Is it possible to derive a method of quantification which will be satisfactory over the *entire* range? To return to the ohms scale in Fig. 1, it will be noted that (*a*) nowhere along the entire scale are the units or intervals of equal size, and (*b*) the size of the intervals increases geometrically with the increase in the general level of resistance. To state this differently, the values of the GSRs along the ordinate may be thought of as being proportional to the relative size of the step intervals which correspond to the numerical values along the abscissa, or the general level of resistance. In fact, this is literally the case, since the average GSR taken at a resistance level of 15,000 ohms covers about one-fifth of the distance between the '15' and '10' positions along the abscissa (since the average GSR here is approximately 1000 ohms decrease in resistance), while the average GSR taken at a resistance level of 50,000 ohms would cover *more* than the distance between the '50' and '45' positions along that portion of the abscissa (since the average GSR here is approximately 6000 ohms decrease in resistance).

If, however, the logarithms (base 10) of the ohms GSRs shown in Fig. 1 are plotted as in Fig. 2, a curvilinear relationship no longer obtains. Rather, we have a straight line, and the slope-intercept definition of the line ($y = ax + k$) becomes: log. ohms GSR = (x, or the slope of the line) $\times$ (resistance level) + k, or the point of intersection of the line with the ordinate. The value of the constant (k) can be subtracted from both sides of the equation (which accounts for its change in sign when the value is added to the log. ohms GSR), and the log. ohms GSR + k value can be divided by the resistance level to leave the constant (x) on the other side of the equation $\left(\text{i.e., } \frac{\text{log. GSR} + k}{\text{resistance level}} = \text{a constant}\right)$. When the actual values shown in Fig. 1 are thus transformed, we find that the new points fall along a line which is approximately horizontal. In the present

set of data, the value of the constant (x) was .0245, which was multiplied by 100 in order to make the units of convenient size.[9]

From the relationship pictured in Fig. 2, it is readily apparent that for verbal stimuli such as were used in the present study, a scale was derived which does possess intervals of approximately equal size. Other things being equal, then, the average GSR of an S with a given resistance level will tend to be comparable to the

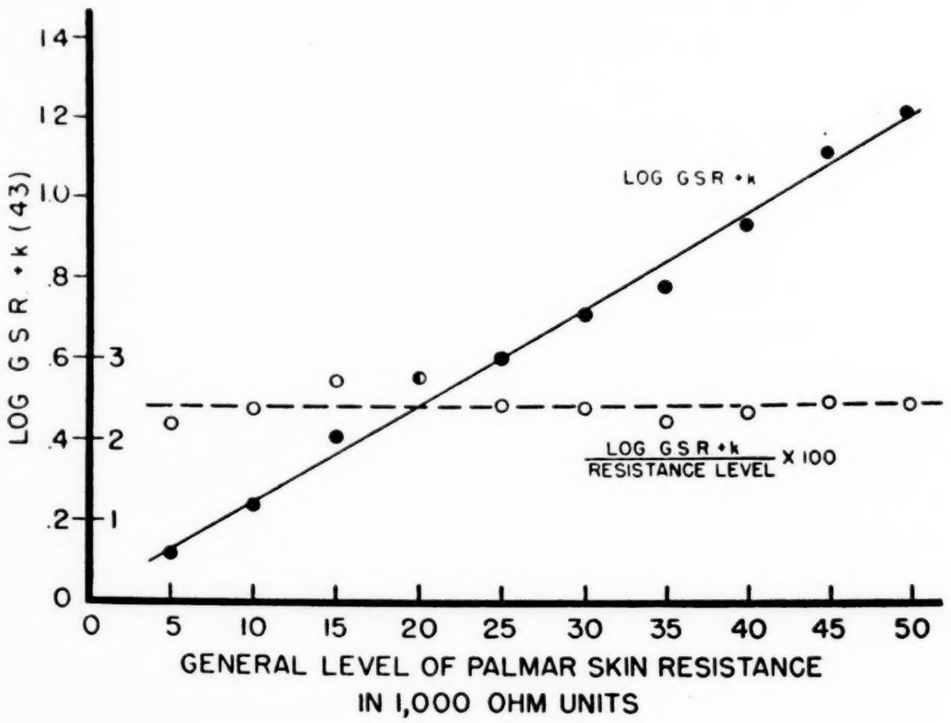

FIG. 2. The size of the log. (ohms) GSR + k units and $\frac{\text{log. (ohms) GSR} + k}{\text{resistance level}} \times 100$ units to 32 verbal stimuli in relation to the general level of resistance. (N = 675, with 48 Ss.)

average GSR of another (or the same) S with a different resistance level if this method of translation is used.[10] However, is the functional relationship shown in Fig. 1 (and which warrants the transformation shown in Fig. 2) limited to verbal stimuli of the type used in this study, or does the same relationship hold for stimuli which are clearly different, whether quantitatively or qualitatively?

[9] It so happens that the ohms GSRs of .1 (or 100 ohms), .2 (or 200 ohms), .3 (or 300 ohms) and .4 (or 400 ohms), if translated into log. GSR + k scores, will result in negative log. values. This difficulty may be overcome by *arbitrarily* assigning the following log. GSR + k, or corrected, values to the respective ohmic values: .1 ohms GSR → .010, log. GSR + k units, .2 → .020, .3 → .030, .4 → .070, .5 → .129. The last value (.5 ohms GSR → .129 corrected GSR) is the same as the one obtained by the proposed formula.

The curves in Figs. 1 to 3 were fitted graphically.

[10] No attempt is made in this paper to relate the GSR values to their psychological or psychophysical correlates.

Supplementary study

In an attempt to find an answer to the question raised above, 380 GSRs to strong electric shock (which were obtained from a total of 66 *Ss*) were analyzed in the same manner as were the 675 GSRs to the verbal stimuli. Although the data for this section were taken from three different experiments, the shocks themselves were of equal strength. These data were combined on the grounds of this equivalence, and are plotted in terms of log. GSR + k units in Fig. 3.

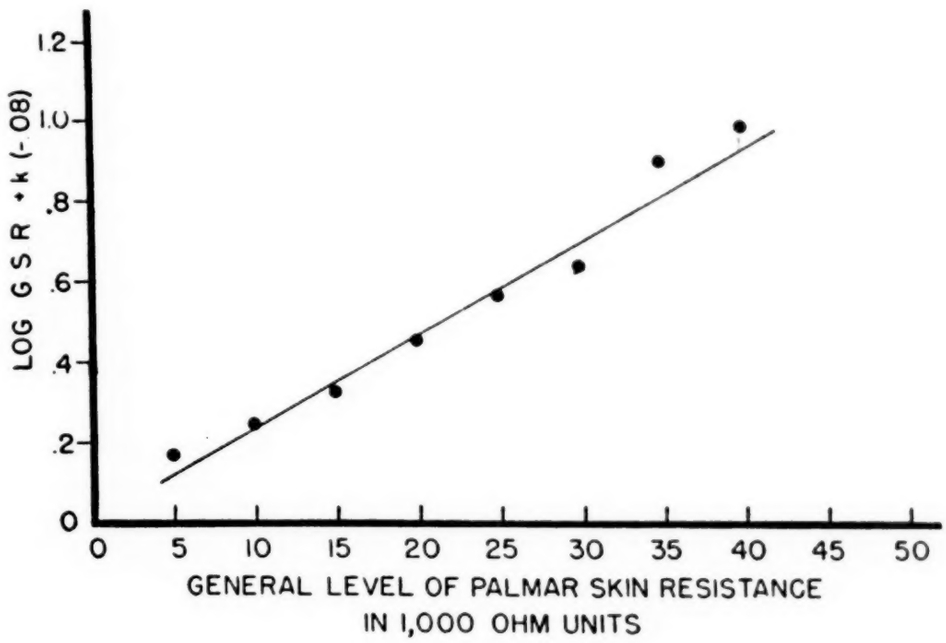

FIG. 3. The size of the log. (ohms) GSR + k units to strong electric shock in relation to the general level of resistance. (N = 380, with 66 *Ss*.)

It is interesting to note that the slopes of the two lines in Figs. 2 and 3 are almost identical, and that the only real discrepancy lies in the different values of the constant (k) which is added to the log. GSR scale. In other words, if the data of Figs. 2 and 3 were plotted along the *same* ordinate (i.e., with the same value of k), the GSRs to the shocks would fall consistently higher at each resistance level than would the GSRs to the verbal stimuli, and the lines drawn through these two sets of points would have approximately the same slope. This finding seems nicely to confirm the assumption that under constant conditions the average size of the drop in skin resistance (i.e., ohms GSR) is a function of the *S*'s general level of resistance, and that the nature of this relationship is approximated by the ohms curve in Fig. 1.

Concluding Remarks

In conclusion, let us evaluate the proposed method of quantifying skin resistance data in view of the four 'requirements' listed earlier in this paper. These requirements or propositions will be considered in turn.

1. *The accuracy with which the transformation can be made.*—The mathematical transformation proposed in this paper is neither so simple nor so convenient as one might desire. This would be particularly true if it were necessary to translate each ohms GSR. But such a prohibitive task can be avoided if a single conversion or equivalence table is constructed which contains the values of most or all of the ohms GSR values for all of the major levels of resistance. Such a table would enable one to determine the corrected values directly, or else make an interpolation which would be satisfactory for all practical purposes.

The question arises, however, as to the universal applicability of such a table. That is, must a separate table be constructed for each experimental situation, or could the same set of values be used for more than one set of data? In attempting to answer this question, it must be remembered that the data pictured in Fig. 1 were recorded by an apparatus which was very accurate, sensitive, linear, etc., and that the *same* relationship might not appear under other conditions —such as when the recording apparatus is non-linear or when the amount of current through the *S* varies with his resistance level.

A further consideration has to do with variations in experimental treatments. It was observed that the value of the constant (k) was clearly different for the verbal ($+ .43$) and the shock ($- .08$) data (e.g., Figs. 2 and 3). But such a state of affairs does not offer any complication if one wishes to make comparisons between the reactions to both the shock and non-shock stimuli within the same experimental framework. In such a case, the value $+ .43$ would be used, and the reactions to shock would be plotted along the ordinate for the non-shock stimuli. But a really satisfactory answer to such questions as the universality of a given set of constants must necessarily await additional research.

2. *Individual differences and experimental treatments.*—In the analysis of individual differences among *S*s, some general method of comparison is necessary. In the present research, the group norms are employed to permit such comparisons. That is, the ultimate numerical value of any particular GSR is relative to the size of the average of all GSRs falling at the same general resistance level. The advantage of equating in terms of the mean GSR value of all *S*s for any given resistance level is that one is then able to make subject-

subject comparisons for the particular resistance level. Similarly, the advantage of equating the mean GSR value in terms of all *Ss* for all resistance levels, as is shown in Fig. 2, is that one is then able to make both subject-subject and treatment-treatment comparisons.[11]

3. *The analysis of general and specific reactions.*—So far only the analysis of the GSRs, or the momentary changes in skin resistance, have been considered. But one may also wish to evaluate the changes in the general level of skin resistance (and the inferred shift in the over-all level of background excitation). In a previous report (**8**), such information proved to be at least as interesting and meaningful as the analysis of the specific changes. But in order to evaluate changes in the general level of resistance, one should first translate the data into the logarithm of the level scores. The reason for using log. level scores is, actually, the same as that for using log. GSR scores—namely that under otherwise comparable conditions the size of the change in ohms resistance increases geometrically with the increase in the level of skin resistance. Thus it is necessary to translate the data in such a manner that the resulting scale will possess units which are roughly equivalent. The multiplication of the raw resistance level scores by their logarithm tends to offset this implicit distortion and to provide an adequate correction. An empirical test of this correction can be made by multiplying the difference between the log. level units (i.e., a given shift in the log. level) by the relative amount of change in ohms resistance occurring over that range. The result will approximate a constant except for the extreme values, and the size of the constant will be proportional to, and hence indicative of, the magnitude of the shift.

4. *The application of statistical tests and intervals of equal size.*—Probably the basic desideratum of any such method of quantifying raw data is that its use should enable the investigator to perform certain mathematical operations. In the application of such statistical techniques as reveal either the probability of a relationship (e.g., C.R., *t*-test, analysis of variance) or the degree of such a relationship (e.g., product-moment or rank-order correlations) which may obtain among the experimental data, one makes the implicit assumption that the 'measuring stick' employed possesses units of equal size. Furthermore, it may be noted that as the size of the units along the scale deviates from uniformity in a systematic manner (as

[11] In the application of the above transformation to the records of individual *Ss*, 84 responses of several *Ss* were plotted as in Figs. 2 and 3. This resulted in two general findings: (*a*) the log. GSR values of any given *S* tend to fall along a straight line, and (*b*) there are variations between *Ss* in the slope of the line and in the point of its intersection with the ordinate (i.e., in the x and k constants mentioned above). Although these variations between *Ss* were not very great, they indicate that at least in certain types of research these measures of individual differences must be considered.

in the ohms scale), the assumption of homogeneity of variance becomes more and more untenable.

From the material presented in this paper, it appears that the proposed method of quantification meets the requirement of equal scale-intervals. That is, for comparable stimuli under comparable experimental conditions, the size of the GSR (or the magnitude of the shift along the measuring scale) is no longer a function of the S's general level of resistance, but rather is constant over the total resistance range.

(Manuscript received July 11, 1944)

Bibliography

1. Darrow, C. W., & Heath, L. L. Reaction tendencies relating to personality. *Studies in the dynamics of behavior* (Ed. K. S. Lashley). Chicago: Univ. Chicago Press, 1932, pp. 59–261.
2. Darrow, C. W. The significance of the galvanic skin reflex in the light of its relation to quantitative measurements of perspiration. *Psychol. Bull.*, 1934, **31**, 697–698. (Abstract)
3. Darrow, C. W. Quantitative records of cutaneous secretory reactions. *J. gen. Psychol.*, 1934, **11**, 445–448.
4. Darrow, C. W. The significance of skin resistance in the light of its relation to the amount of perspiration. (Preliminary note) *J. gen. Psychol.*, 1934, **11**, 451–452.
5. Darrow, C. W. The equation of the galvanic skin reflex curve: I. The dynamics of reaction in relation to excitation-background. *J. gen. Psychol.*, 1937, **16**, 285–309.
6. Davis, R. C. Factors affecting the galvanic reflex. *Arch. Psychol.*, 1930, No. 115, 1–64.
7. Freeman, G. L., & Katzoff, E. T. Methodological evaluation of the galvanic skin response, with specific reference to the formula for R. Q. (recovery quotient). *J. exp. Psychol.*, 1942, **31**, 239–248.
8. Haggard, E. A. Experimental studies in affective processes: I. Some effects of cognitive structure .and active participation on certain autonomic reactions during and following experimentally induced stress. *J. exp. Psychol.*, 1943, **33**, 257–284.
9. Haggard, E A. An apparatus for the measurement of changes in palmar skin resistance. (In preparation)
10. Hunt, W. A., & Hunt, E. B. A comparison of five methods of scoring the galvanic skin response. *J. exp. Psychol.*, 1935, **18**, 383–387.
11. Wenger, M. A., & Irwin, O. C. Fluctuations in skin resistance of infants and adults and their relation to muscular processes. *Studies in infant behavior*, III, *Univ. Ia. Stud. Child Welf.*, 1936, **12**, No. 1, pp. 143–179.

A STUDY OF INDIVIDUAL DIFFERENCES IN MOTION ACUITY AT SCOTOPIC LEVELS OF ILLUMINATION *

BY C. J. WARDEN, H. C. BROWN AND SHERMAN ROSS [1]

Columbia University

Introduction

Little is known regarding motion acuity at parafoveal and peripheral retinal positions under low levels of illumination. The same may be said of the relationship between form and motion acuity at both photopic and scotopic levels of illumination. The purpose of the present investigation was to study certain aspects of these relationships.

This paper is a continuation of the work recently reported bearing on the relationship between form and motion acuity (**2**). The earlier report was concerned mainly with the standardization of apparatus and method for use in the present investigation. The emphasis throughout the present study is upon the factor of individual differences.

From the standpoint of military psychology, the primary interest was in the selection of night-fliers. The question had been raised as to whether or not the Snellen Test is adequate for this purpose. If not, the problem of developing a test of scotopic acuity would need to be solved. Moreover, it seemed to the present authors that motion acuity rather than form acuity ought to be stressed in the field of aviation. For motion acuity is obviously the most important visual function utilized in flying. In fact, the logic of the situation would suggest the use of a scotopic motion acuity test as the most appropriate screening device.

The specific attack upon the problem was made along the following lines in the present study. First of all, a group of *S*s was selected on the basis of Snellen scores. These *S*s were then dark-adapted for an appropriate period, and given the Motion Acuity Test at two low levels of illumination. Threshold determinations were made through a series of six retinal positions, ranging from 7 degrees to 55 degrees of peripheral angle.

* The work described in this paper was done under a contract recommended by the Committee on Medical Research, between the Office of Scientific Research and Development and Columbia University.

[1] The senior author was Responsible Investigator, and the junior authors were Professional Scientific Assistants, on this project (OEM cmr-264).

This experimental design enabled us to make such comparisons as the following: (1) the range of Snellen Scores and of the motion acuity threshold values of the group, (2) the correlation between Snellen Scores and motion acuity threshold values, (3) the comparison of the motion acuity threshold values for the two levels of illumination tested, and (4) the application of these findings to hypothetical combat conditions.

Method and Procedure

In this section, a brief description of the *S*s, the apparatus, and the test procedure is presented.

Snellen test.—This test was administered under standard conditions of distance and illumination to each *S* prior to acceptance. A group of 28 male students at Columbia University were selected by this means. They were between the ages of 17 and 26 years, and had a Snellen Index of 20/20 or better in both eyes, without glasses, and reported no history of gross visual defect. Most of the men had already passed the visual requirements for the Army or Navy Officer's Reserve. They were retested, however, before being accepted as *S*s. The actual Snellen Scores of those accepted were recorded for use in computing correlations between photopic form acuity and scotopic motion acuity. The test of motion acuity required three sessions of about two hours each in the dark-room. The *S*s were paid a nominal sum at the close of the experiment, and were well motivated throughout.

Motion acuity apparatus.—The apparatus has been described in detail in a previous paper (2). It consisted essentially of the following parts: (*a*) A motor-gear-pulley system for rotating the stimulus. This system could be manipulated by the *S* in securing a threshold speed (method of adjustment). (*b*) A dark cubicle for the *S*, lighted from above by a projector set for a given level of illumination. A relatively high (log. −3.4 ml.) and a relatively low (log. −4.0 ml.) scotopic level of illumination was used. The latter is presumably about the lowest level involved in ordinary night-flying. (*c*) A perimetry set-up for securing the desired peripheral angle. A series of six retinal positions was employed: 7, 10, 25, 35, 45, and 55 degrees from the fovea. (*d*) A series of stimuli, in various sizes, as described in detail in Table I. The stimulus figure consisted

TABLE I

Description of Standard Stimulus Set

Peripheral Angle at Which Stimulus was Used	Circumference in Millimeters	Arc of Gap in Millimeters	Arc of Gap in Degrees of Visual Angle
7, 10	160.3	8.9	1.04
25	188.6	10.5	1.22
35	220.0	12.2	1.43
45	285.1	15.8	1.85
55	349.0	19.4	2.25

The manner of determining the size of this series of stimuli has been discussed in an earlier paper (2). They were large enough for the *S*s with poorer motion acuity, but larger than necessary for the better *S*s.

of a Landolt ring modified for rotary motion. Instead of a single gap, there were nine parallel-sided breaks or gaps (white areas), equidistant from each other. A white bristol board background, bent in the proper arc for perimetry, served as the surround. The stimulus was exposed by means of a white bristol board shutter operated manually.

As will be noted, the stimulus involved a circular figure of black wedges against a white surround and general background. At one stage in the experiment, these brightness relationships were exactly reversed to see whether or not such a change would make any difference in motion perception. Since this stimulus-background change was not effective in raising or lower-

ing the motion threshold, no further mention of this variation from the standard procedure will be made.

Procedure.—Motion threshold determinations were made on the right eye only, the left being covered with an Adjusto eye-shield. This device permitted the eye to remain open, to blink and wink, but prevented binocular accommodation on the fixation point.

Each *S* usually required three sessions of about three hours each, given in the following order: (*a*) training session, (*b*) experimental session at the log. −4.0 ml. level, and (*c*) experimental session at the log. −3.4 ml. level. Each session was preceded by a dark-adaptation period of 25 min. (Navy red dark adaptor goggles), followed by a 5-min. period within the cubicle, in order to reach the precise testing level.

The *S*s were all untrained in this type of observation, hence the first session was devoted largely to training and demonstrations. Each *S* was given information concerning the anatomy of the eye, the specific functions of the rods and cones, the meaning of the curve of dark adaptation, etc. He was then instructed how to fixate, how to operate the rheostat controlling the speed of the stimulus, how to relax between judgments, etc. He was also trained in the use of the criterion of threshold motion, and was required to make sample judgments. During this session, he became adjusted to a three-second exposure time. The *S*s were usually proficient in making threshold judgments by the end of the first session.

In the two test sessions, the method of adjustment was applied as follows in determining the threshold rate of movement: (1) the *E* set the general speed level far above the threshold speed, (2) the *S*, using the rheostat control, then decreased the speed in large steps to a rate below the threshold, (3) the *S* then increased the speed in small steps up to the threshold level, in terms of our criterion.

A high criterion was used here in order to avoid doubtful judgments and inconsistencies. At the threshold level, the *S*s were required to see definite motion with absolute certainty. If the motion was so slow as to appear doubtful, or to involve inference or suggestion, it was not regarded as criterial. In such a case the *S* was required to establish the threshold value at a somewhat higher speed. The standard requirement was that the *S* be able to make a judgment of good motion in two out of three exposures at the same speed. Three determinations were made for each retinal position at both levels of illumination. The threshold values determined by the use of this criterion were probably somewhat above the absolute threshold.

The exposure time was controlled by the shutter. When the *S* indicated that he was 'ready,' the *E* counted aloud '1–2–3' and raised the shutter. At the end of three seconds, the shutter was closed. A three-second exposure time had been found to be adequate in preliminary experiments (2).

The threshold speed of rotation of the stimulus was read from the chronoscope, in terms of seconds, for a given angle on the timing disc. These values were then transmuted into degrees of visual angle per second, as the most convenient index of motion acuity.

Results

The results secured on the 28 *S*s for both levels of illumination are given in Tables II and III, and the data are plotted in Figs. 1 and 2. The threshold values for motion acuity are given in terms of degree of visual angle per second. These measures represent the threshold rates of rotation of the stimulus at the various retinal positions, under each level of illumination.

The general form of the function is essentially the same for both levels of illumination. However, the curves for the higher level (log. − 3.4 ml.) lie below the corresponding curves for the lower level (log. − 4.0 ml.). This relationship was found to be consistently true for all *S*s, as might have been expected.

In general the curves show a rise in trend as the peripheral angle is increased. This trend is very definite from the 10 degree retinal

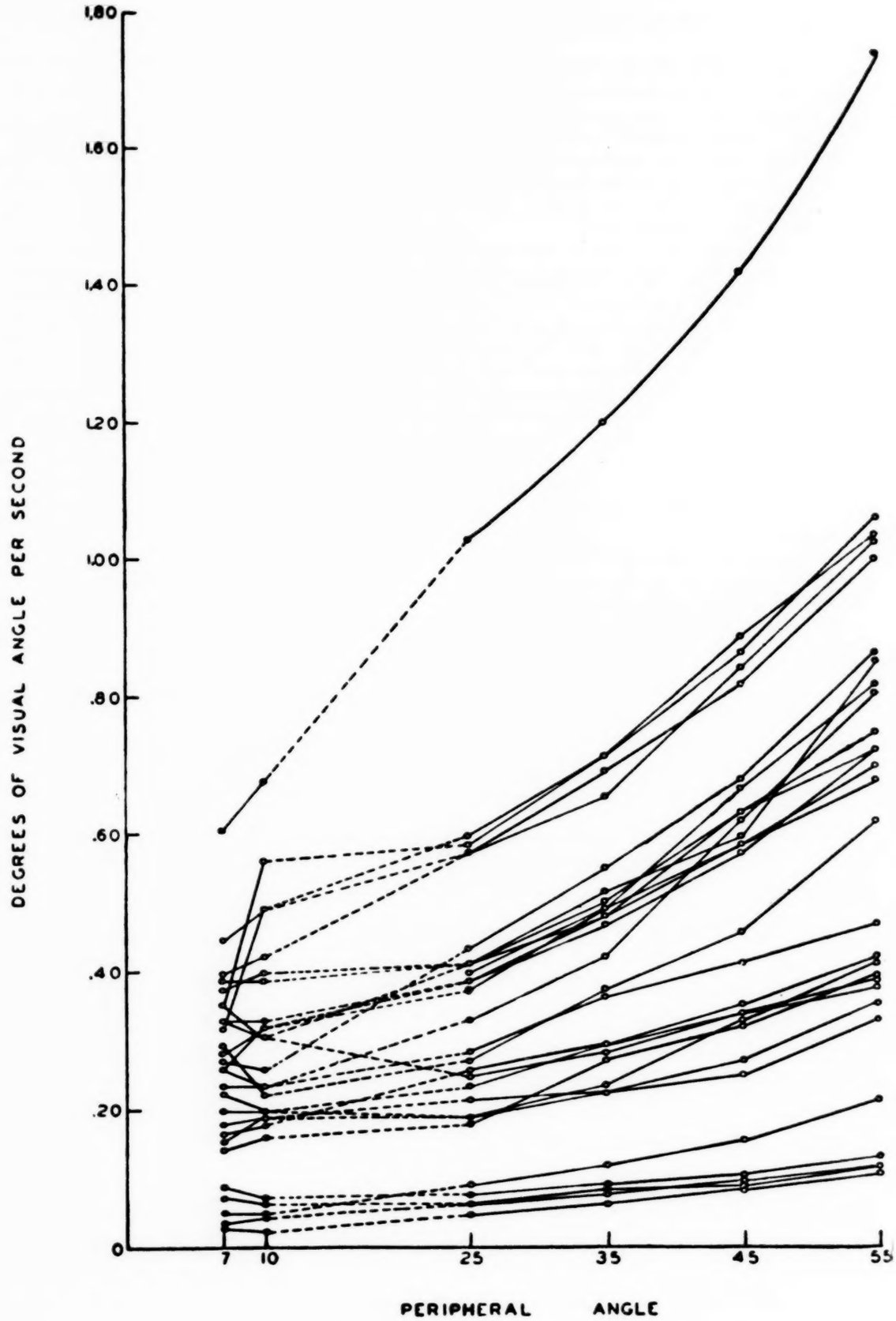

FIG. 1. Log. −3.4 ml. level of illumination.

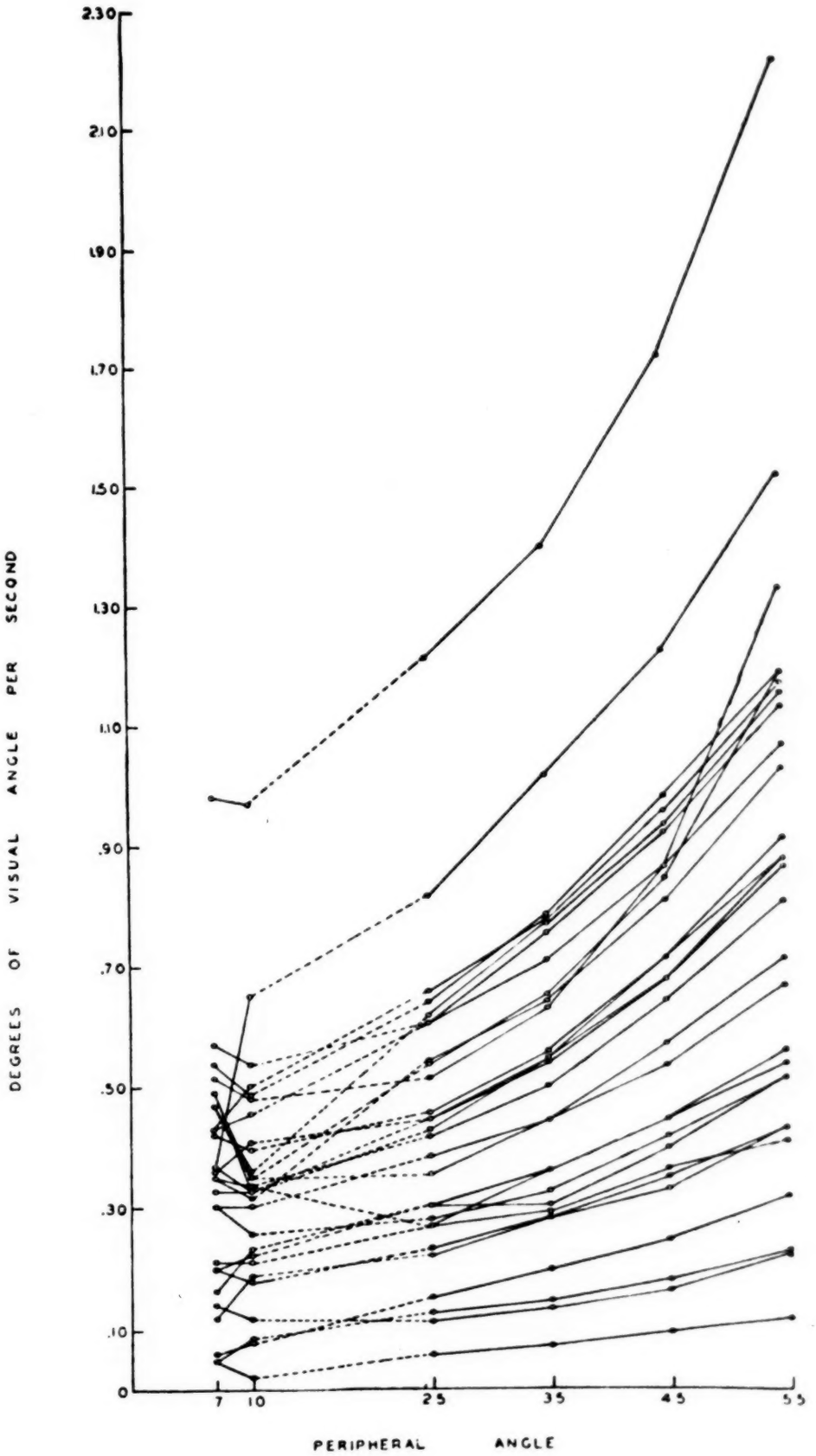

FIG. 2. Log. −4.0 ml. level of illumination.

position outward. The results for the 7 degree position are somewhat irregular. This was due to the fact that the rather large stimulus used here was very close to the fixation point, and overlapped the near-foveal region. The data at this point may be ignored, however, since the 10 degree position is commonly regarded as the most sensitive (1).

When the threshold values are plotted on semi-log paper (Figs. 3 and 4), a straight line relationship obtains. This is especially clear

TABLE II

RESULTS FOR THE HIGHER SCOTOPIC LEVEL OF ILLUMINATION (LOG. −3.4 ML.) IN TERMS OF DEGREES OF VISUAL ANGLE PER SECOND

Subject	Retinal Position					
	7°	10°	25°	35°	45°	55°
1	.608	.679	1.030	1.205	1.427	1.755
2	.351	.562	.585	.714	.889	1.041
3	.293	.222	.269	.374	.456	.620
4	.234	.234	.328	.421	.620	.807
5	.257	.316	.386	.480	.632	.749
6	.386	.386	.410	.515	.597	.854
7	.269	.257	.433	.550	.679	.866
8	.328	.304	.410	.480	.585	.702
9	.281	.316	.374	.491	.667	.819
10	.374	.398	.410	.503	.632	.725
11	.351	.304	.398	.491	.585	.679
12	.328	.328	.386	.468	.573	.725
13	.199	.199	.234	.293	.351	.421
14	.257	.234	.281	.363	.410	.468
15	.222	.199	.187	.222	.269	.351
16	.164	.176	.257	.293	.339	.386
17	.176	.187	.211	.222	.246	.328
18	.152	.187	.187	.234	.328	.410
19	.087	.071	.075	.089	.103	.129
20	.140	.158	.176	.269	.319	.391
21	.328	.304	.246	.281	.339	.374
22	.049	.049	.090	.117	.152	.211
23	.034	.043	.063	.076	.095	.113
24	.071	.062	.062	.074	.089	.112
25	.026	.020	.047	.061	.082	.105
26	.316	.491	.597	.714	.866	1.065
27	.445	.491	.573	.690	.819	1.006
28	.398	.421	.573	.655	.842	1.030

from the 25 degree retinal position outward. The curves, in general, take the form of an exponential function, such as $y = ka^x$.

The range of individual differences in motion acuity threshold values is extremely marked. These values for the 10 degree retinal position at the lower brightness level (log. − 4.0 ml.) are plotted in Fig. 5. As will be seen, the histogram takes the general form of a normal distribution with some tendency toward positive skewness. It seems likely that with a sufficiently large number of cases a fairly

typical normal curve would be obtained. This would mean that a group, highly selected on a cone level form test (Snellen), would be practically unselected so far as scotopic motion acuity is concerned. In brief, the fact of normal (20/20 or better) photopic form vision offers no basis for predicting high motion acuity at low levels of illumination.

This conclusion is further substantiated by a correlation between Snellen Index and motion acuity threshold value. The biserial cor-

TABLE III

RESULTS FOR THE LOWER SCOTOPIC LEVEL OF ILLUMINATION (LOG. −4.0 ML.) IN TERMS OF DEGREES OF VISUAL ANGLE PER SECOND

Subject	Retinal Position					
	7°	10°	25°	35°	45°	55°
1	.983	.971	1.217	1.404	1.720	2.223
2	.374	.655	.819	1.018	1.229	1.521
3	.515	.480	.515	.632	.866	1.322
4	.468	.351	.538	.655	.842	1.193
5	.468	.363	.620	.772	.936	1.158
6	.573	.538	.608	.761	.924	1.135
7	.351	.316	.542	.644	.819	1.030
8	.421	.398	.456	.562	.714	.913
9	.491	.328	.445	.550	.714	.878
10	.363	.410	.445	.538	.679	.878
11	.374	.328	.433	.550	.679	.866
12	.351	.339	.421	.503	.644	.807
13	.304	.304	.386	.445	.573	.714
14	.491	.351	.363	.445	.538	.667
15	.211	.211	.269	.363	.445	.562
16	.164	.234	.304	.363	.445	.538
17	.304	.257	.281	.328	.421	.515
18	.199	.222	.304	.304	.398	.515
19	.117	.187	.222	.281	.351	.433
20	.199	.176	.234	.281	.328	.433
21	.339	.339	.269	.293	.363	.410
22	.059	.077	.152	.199	.246	.316
23	.046	.082	.129	.146	.181	.228
24	.140	.117	.111	.129	.164	.222
25	.047	.023	.056	.070	.094	.117
26	.538	.491	.644	.784	.983	1.193
27	.433	.503	.661	.778	.959	1.170
28	.433	.456	.610	.714	.866	1.077

relation technique was used. This involved arranging the individuals into two Snellen categories (20/20, 20/15 or better), and ranking the mean motion acuity threshold values for all six retinal positions at a given level of illumination. The biserial correlation for the higher level of illumination (log. − 3.4 ml.) was found to be only .07 ± .10. This means that an individual with a high Snellen Index (20/15 or better) is not likely to have any better scotopic motion acuity than an individual with an average Snellen Index (20/20).

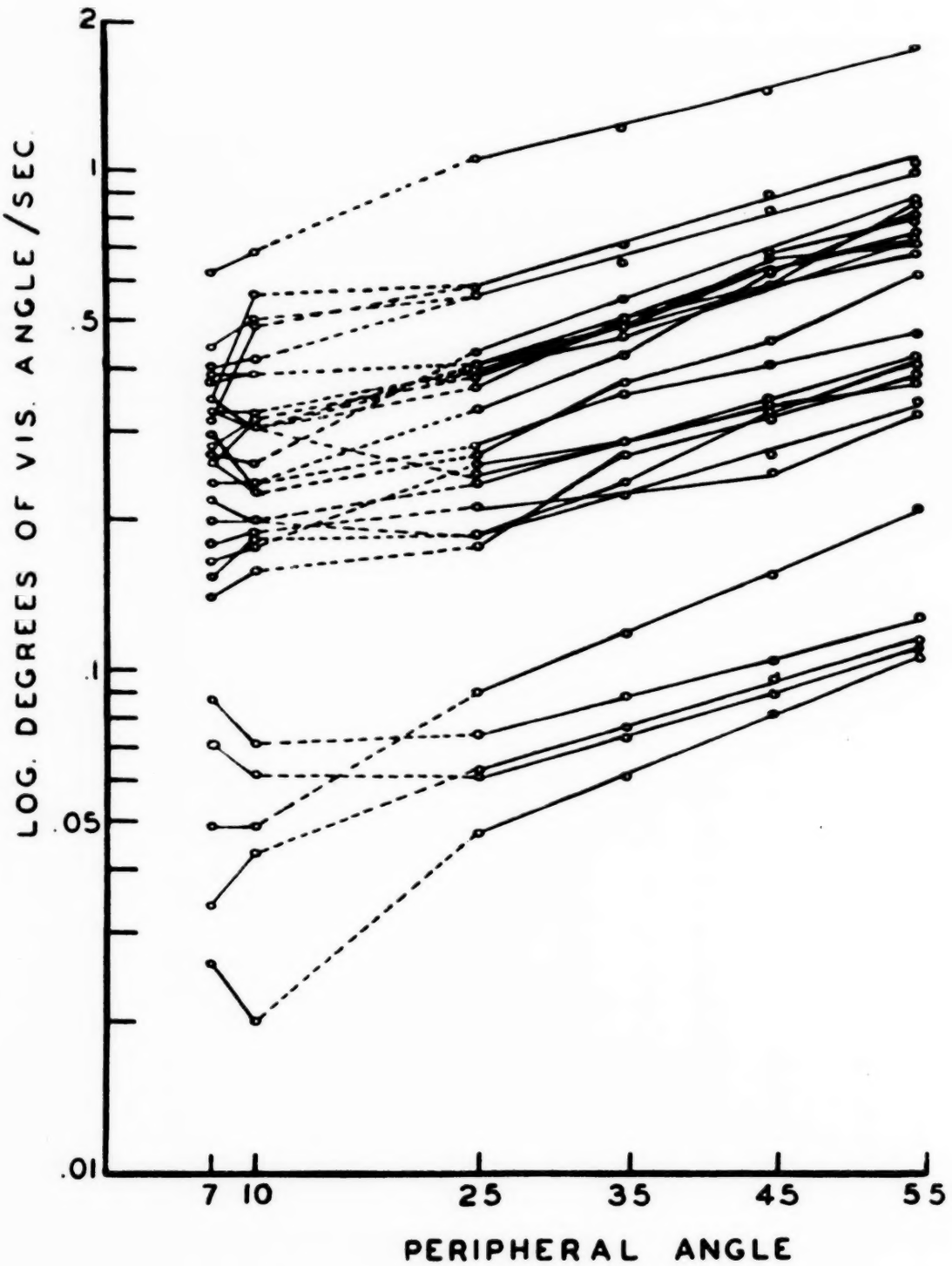

FIG. 3. Semi-log curves (log. −3.4 ml.).

APPLICATION TO NIGHT-FLYING CONDITIONS

In the present section an attempt is made to apply our results to the night-flying situation. The question of major interest is as to whether or not the individual differences in motion acuity here reported are large enough to influence efficiency in night-flying.

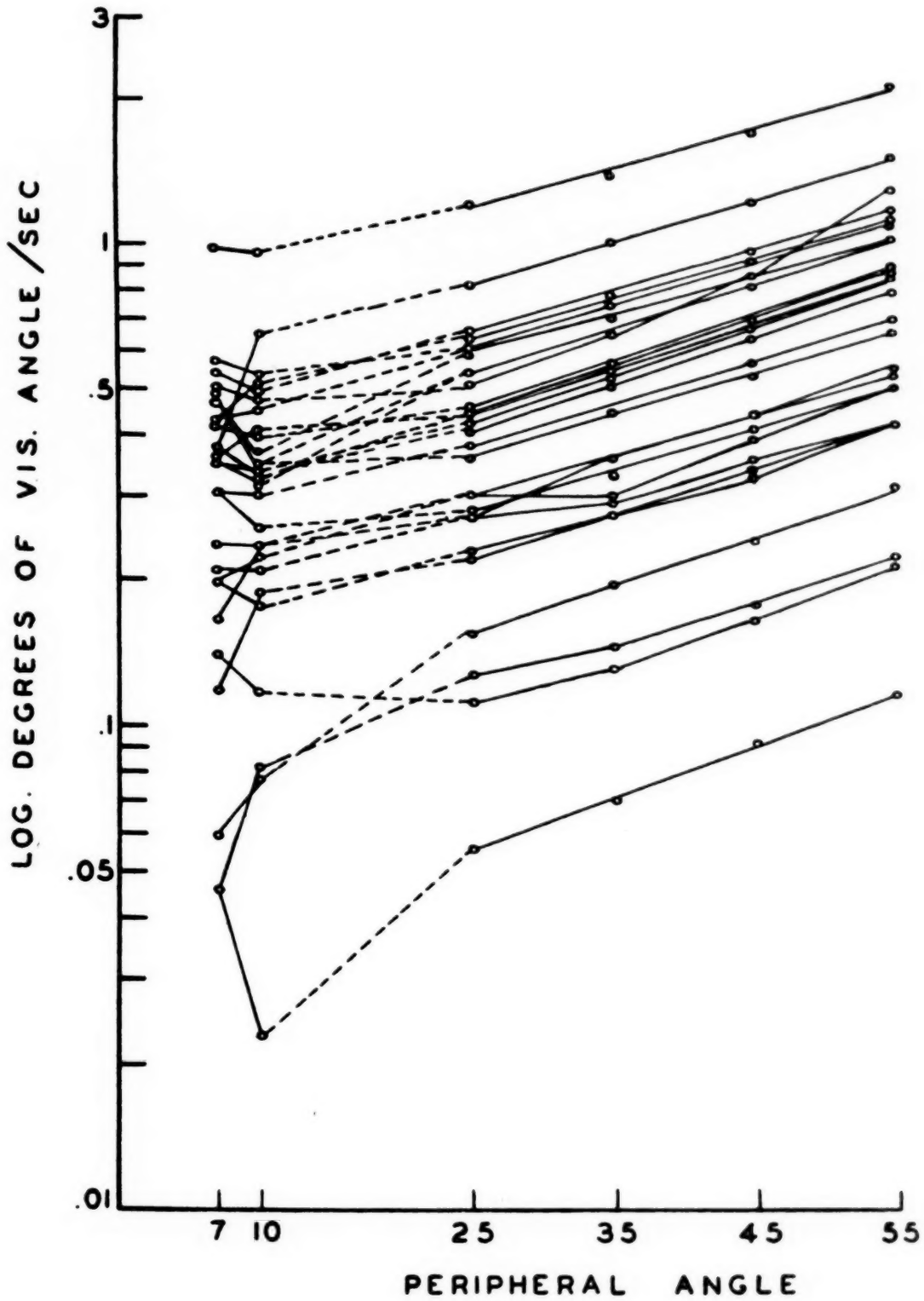

FIG. 4. Semi-log curves (log. −4.0 ml.).

Are some of our Ss so insensitive to moving objects at low illuminations that they would not make good night-fliers?

We found it impossible to obtain precise information on the actual flying situation. It was necessary, therefore, for us to base our computations upon a set of hypothetical conditions, which would seem to agree fairly well with the actual flying situation. In the computations reported, the specific conditions assumed are as follows: (1) size of plane: 90 feet in the longest dimension (wingspread), (2) distance of this object from the S: 1.5 or 2.0 miles, and (3) level of illumination: scotopic (log. − 4.0 ml.).

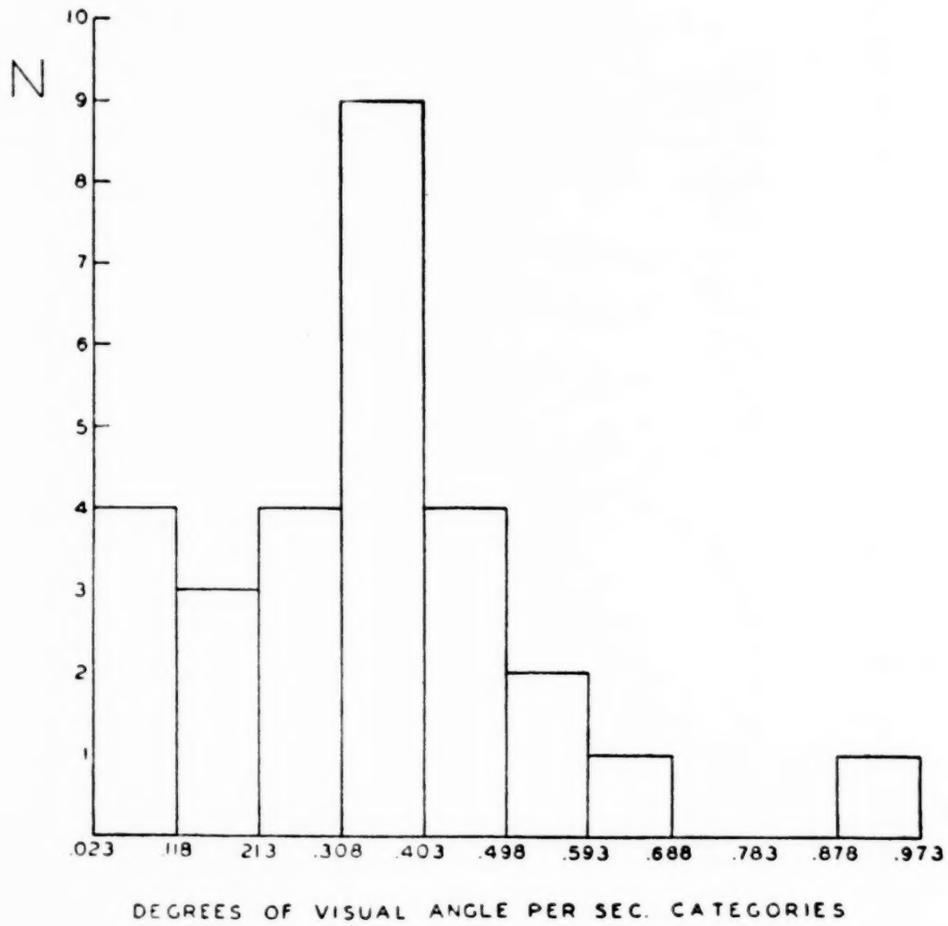

FIG. 5. Frequency distribution of motion acuity threshold values at the 10 degree retinal position (log. −4.0 ml. level of illumination).

There is reason to believe that the distances indicated above represent about the limit of form visibility for a plane of this size under this low level of illumination. These values (1.5 and 2.0 miles) correspond closely with the projection of the form threshold value, determined on four Ss, at this level of illumination. These form thresholds represent a reduction in visibility to about 1/20 of the

normal distance perception in daylight, as might be expected. As a matter of fact, the stimuli used in determining the motion acuity thresholds were somewhat larger than the stimuli used to secure the form acuity thresholds for projection. It seems reasonable to assume, therefore, that a plane of this size would be definitely visible at 1.5 or 2.0 miles.

Individual thresholds in motion acuity were computed to agree with the specific setting indicated above. The threshold values at the 10 degree retinal position (log. − 4.0 ml.) were taken as the basis

TABLE IV

INDIVIDUAL DIFFERENCES IN MINIMUM SPEED NECESSARY FOR THE OBSERVATION OF A PLANE (90 FEET WINGSPREAD) AS MOVING

Subjects	Threshold in Degrees of Visual Angle/Sec.	Speed in Miles per Hour at Which Motion is First Seen	
Rank Order		Plane 1.5 Mi. Distant	Plane 2 Mi. Distant
1	.023	2.15	2.86
2	.077	7.21	9.59
3	.082	7.68	10.21
4	.117	10.96	14.58
5	.176	16.49	21.93
6	.187	17.52	23.30
7	.211	19.76	26.28
8	.222	20.79	27.65
9	.234	21.92	29.15
10	.257	24.07	32.01
11	.304	28.48	37.88
12	.316	29.60	39.37
13	.328	30.72	40.86
14	.328	30.72	40.86
15	.339	31.75	42.23
16	.339	31.75	42.23
17	.351	32.88	43.73
18	.351	32.88	43.73
19	.363	34.00	45.22
20	.398	37.73	49.78
21	.410	38.40	51.07
22	.456	42.71	56.80
23	.480	44.96	59.80
24	.491	45.99	61.17
25	.503	47.12	62.67
26	.538	50.39	67.02
27	.655	61.35	81.60
28	.971	90.95	120.96

Note.—Ss are ranked in order of threshold motion perception at 10° peripheral angle at level log. −4.0 ml. The speed indicated in the last two columns refers to the difference in speed (resultant) between the observer's plane and the observed plane.

of computation, since this retinal position is generally regarded as the most sensitive at low illumination. The results of this analysis are presented in Table IV. The threshold values obtained in our experiment, arranged in rank order, are given in column 2 of the

table. The values in the last two columns of the table represent the threshold rate in miles per hour for the perception of motion of a 90-foot plane at the distances indicated. These speed values refer to the *resultant* speed (difference) between the plane of the observer and the plane to be observed.

The meaning of the speed values in the table can best be indicated by a few examples. The best *S* (No. 1) should be able to detect a plane as moving when the resultant speed is only 2.15 miles per hour at a distance of 1.5 miles. The rate for a distance of two miles would need to be 2.86 miles per hour. On the other hand, the worst *S* (No. 28) could not detect the plane as moving, unless the resultant speed were 90.95 miles per hour at 1.5 miles, or 120.96 miles per hour at a distance of 2.0 miles. The individual differences in the rate of movement of the object, for the threshold perception of motion under these conditions, are thus seen to be extremely large.

In order to determine whether or not our poorer *S*s are too insensitive to moving objects to make good night-fliers, the critical resultant speeds for actual combat should be known. We have been unable to secure definite information on this point. The demands for the perception of movement at different rates would vary from one type of plane to another, and from one combat situation to another.

However, it is possible to demonstrate how the information in Table IV might be applied to the selection of night-fliers. This can be done by assuming arbitrarily that certain resultant speeds are critical. For example, if the resultant speed of 50 miles per hour is critical, *S*s 26, 27, and 28 would be unable to see the plane as moving, at a distance of 1.5 miles. At a distance of two miles, *S*s No. 21 through 28 would not be able to observe it as moving. If we assume a resultant speed of 75 miles per hour to be critical, then *S* 28 would be rejected at the 1.5 mile distance, and *S*s 27 and 28 at the 2.0 mile distance. Even at this high critical speed, 3.6 percent of the *S*s would be eliminated at the shorter distance, and 7.2 percent at the longer distance.

A Short Test for Night-Fliers

Although the present experiment involved the use of complicated apparatus, a much-simplified test for practical use in the selection of night-fliers could be readily developed. Such a standard test would require only a short time per *S*, and would be easy to administer. On the basis of the findings of the present experiment, a test for selection could be simplified along the following lines:

1. The apparatus for rotating the stimulus would be small, com-

pact and portable. By adapting the method used in the Navy Adaptometer (3) to motion acuity, only a single speed would be required. This would be the median threshold speed, as determined on a group of 100 or more *Ss*. Individual ratings would be based upon the percentage of correct responses to this median speed on 50 or more trials.

2. It would be restricted to a single rod level of illumination. This is possible since a high correlation (.94) was found between threshold values at the two low brightness levels used in the present experiment.

3. Binocular vision would be used, with a fixation point approximately 10 degrees above the stimulus (3). The determinations made at a single retinal position would serve as a dependable index of scotopic motion acuity. This is true since the motion acuity curves obtained in the present study were regular out to 55 degrees.

4. It is estimated that the time required to give a screening test with such an apparatus would be about 10 min. per *S*, in addition to the time required for dark adaptation.

Summary

A group of 28 male *Ss*, between 17 and 26 years of age, with a Snellen Index of 20/20 or better were tested for motion acuity. Threshold values were secured under two levels of illumination (log. − 3.4 ml. and log. − 4.0 ml.), at six retinal positions ranging from 7 degrees to 55 degrees.

The following facts and conclusions emerge from the experiment:

1. The motion acuity function for each level of illumination was found to be regular, with a gradual rise as the peripheral angle increased. As might be expected, the curves for the higher level (log. − 3.4 ml.) were much lower than the curves for the lower level (log. − 4.0 ml.). The rank order correlation between the average motion acuity threshold values for the two levels is .94 ± .18.

2. The range of individual differences in motion acuity threshold values is extremely marked, and shows a steady increase from the 7 degree retinal position outward. At the 10 degree retinal position (log. − 4.0 ml. level) the range in threshold values is from .023 to .971 degrees of visual angle per second (Table III). The threshold value of the worst *S* is thus over 40 times as large as that of the best *S*. An attempt to apply these results to night-flying conditions is discussed in a separate section.

3. No significant relationship was found between the Snellen Index and the motion acuity threshold value. A biserial correlation was computed between two Snellen categories (20/20 and 20/15 or

better) and the motion acuity threshold values at the log. − 3.4 ml. level of illumination. This biserial correlation was only .07 ± .10. A Snellen Index of 20/20 is no guarantee that an aviator possesses adequate motion acuity at scotopic levels of illumination.

4. A plan for a short and simplified test of scotopic motion acuity, to be used in the selection of night-fliers, is outlined in the final section of this report.

(Manuscript received May 9, 1944)

Bibliography

1. Liljencrantz, E., Swanson, C. A., & Carson, L. D. The use of the eyes at night. *Proc. U. S. Nav. Inst.*, 1942, **68**, 802–810.
2. Warden, C. J., & Brown, H. C. A preliminary investigation of form and motion acuity at low levels of illumination. *J. exp. Psychol.*, 1944, **34**, 437–449.
3. Webster, A. P. The measurement of thresholds of perception and discrimination. 1943. Pp. 18. (Unpublished)

GENERALIZATION IN MEMORY [1]

BY JEAN GOODWIN, LOUIS LONG AND LIVINGSTON WELCH

Hunter College and College of the City of New York

It has been clearly demonstrated in the field of conditioning that if a conditioned stimulus is established to a given response, other stimuli in the same modality will elicit the same response without added reinforcement (**7, 8**). These studies have further suggested that the magnitude of the generalized response varies inversely with the difference between the related stimulus and the conditioned stimulus.

Cofer and Foley (**1, 2**) have used this principle of generalization in interpreting verbal behavior. For example, generalization in the verbal sphere may be mediated semantically (this is, in terms of related meanings of the stimulus and the reinforcing word) and homophonically (that is, in terms of the related auditory qualities of the stimulus and the reinforcing word). In their experiment (**2**) one group of subjects was presented with synonyms of the words to be recalled later, another group was presented with homophones of the words to be recalled later, and a third group was presented with words bearing no relation to the words to be recalled later. All three groups were presented with the same recall list. The recall scores for the first two groups were significantly greater than that for the third group.

In the present experiment the writers were interested in determining the effect of still another type of verbal reinforcement upon the mediation of generalization. The relationship chosen for study is a semantic one and involves the use of concepts at different hierarchy levels. The stimulus word list was composed of generic terms, while the reinforcing list contained words which were members of the genus classes presented in the stimulus word list (e.g., color and red, members of the stimulus and reinforcing lists, respectively). The primary purpose of the experiment was to determine whether this species-genus relationship, so essential in reasoning, would have an effect on the recall of stimulus words provided the subjects were trained and tested for recall on words bearing this relationship to one another.

[1] From the Institute for Research in Child Psychology, Hunter College. The authors are indebted to Miss R. M. Thomas for her assistance in collecting part of the data of this experiment.

Method and Procedure

Word lists.—The words used throughout the experiment were selected from the 2000 most frequently used words as determined by Thorndike (9); they are listed in Table I.

TABLE I

Stimulus Words Used in Experiment

Practice List	Equating List	Specific Object List	Control List	First Hierarchy List *
port	Indian	1 dog	1 ring	animal
see	land	2 red	2 teacher	color
toy	age	3 date	3 book	fruit
money	angel	4 tree	4 door	plant
sun	bee	5 five	5 day	number
pen	daughter	6 supper	6 needle	meal
water	desk	7 coat	7 fork	clothes
	Tuesday	8 bean	8 diamond	vegetable
	rain	9 room	9 minister	house
	owl	10 copper	10 sailor	metal
	moon			flower
	lady			parent
	nurse			tool
	class			army
	art			cattle

* The first 10 words of this list are *test words* (i.e., the genus name for words of the Specific Object List); the last five words are *control words* which have no relationship to the words of the Specific Object List.

Subjects.—Two groups of children from the Hunter Elementary School were selected as *Ss*. The age range of the *Ss* was from 8 to 10 years. The 27 members of each group were matched on the Equating List (Table I) in terms of the number of words recalled after one presentation of the list.

Apparatus.—The Gerbrands Memory Drum was used to present the stimuli.[2] Each list was typed on a white paper band which was slipped onto the drum of the apparatus. The words were exposed singly for one and one-half sec. An experimental session required from 10 to 15 min.

Testing procedure.—(*a*) *Preliminary practice.*—Each *S* was instructed as follows: "Watch this opening (*E* pointing to the window in the shield of the apparatus) carefully. You will see a list of words presented one at a time. As each word is presented *say it aloud and think about what it means.* When all of the words have been shown, I am going to ask you to tell me as many of the words as you can remember. Are there any questions?" The *S* was then shown the Practice List of Table I. During the recall period *E* recorded the words as they were recalled. (*b*) *Equating List.*—Without further instructions *S* was shown the words in the Equating List (Table I), and when all of the words had been presented *E* recorded the words recalled by *S*. On the basis of performance on this list the *Ss* were divided into two groups (experimental and control) so that the mean number of words recalled was the same for both groups (mean = 5.8). (*c*) *Specific Object List.*—The instructions of (*a*) were repeated except that the section on recall was omitted. The experimental group was then shown the Specific Object List and the control group was shown the Control List. The lists were presented four times in succession to the respective groups. The random order of the words in these successive trials was as follows (the numbers refer to the number which is placed in front of each word of these two lists in Table I):

Trial 1: 10, 5, 1, 7, 6, 8, 3, 4, 9, 2;
Trial 2: 5, 9, 2, 10, 7, 3, 1, 8, 6, 4;
Trial 3: 9, 3, 1, 4, 5, 6, 7, 10, 2, 8;
Trial 4: 1, 5, 9, 7, 10, 8, 6, 3, 4, 2.

[2] The authors would like to take this opportunity to express their thanks to the members of the Department of Psychology of the College of the City of New York for permitting us to use the department's memory drum.

(*d*) *First Hierarchy List.*—This list was presented to all *S*s. They were reminded about pronouncing the words as presented and were told that recall would take place after one presentation of the list. *E* recorded the words as they were recalled by the *S*s. The words presented are listed in Table I. It should be noted that the first 10 words are generic terms of the 10 words of the Specific Object List (referred to hereafter as *test words*), while the last five words of the list are unrelated to the words of the Specific Object List (referred to hereafter as *control words*).
(*e*) *Introspections.*—Upon the completion of the testing all of the *S*s were asked the following questions:

(1) Do you have anything you would like to say about the words?
(2) Did anything in particular help you in recalling the words? Did you use any special aids?
(3) Did you notice any relationship between the words on any of the lists? If so, what was the relationship?

Results and Discussion

An analysis of the data in terms of the number of words recalled reveals that the mean number of *test words* of the First Hierarchy List recalled by the experimental group was 5.41, while the mean number of words recalled by the control group was 3.22. Using Student's *t*-test the difference between the means was found to be statistically reliable (see Table II). However, when the two groups

TABLE II

Comparison of the Experimental and Control Groups on the First Hierarchy List

Test words

Groups	No. of Subjects	Mean No. of Words Recalled	S.D.	Student's *t*	*P*-value
Experimental	27	5.41	1.65	5.21	<0.01
Control	27	3.22	1.23		
Control words					
Experimental	27	2.07	0.95	0.70	>0.40
Control	27	2.26	0.96		

were compared on the basis of the recall score for the *control words* of the First Hierarchy List a reliable difference was not found between the mean scores (Table II).

One factor which might have been influential in determining the amount of recall was that of practice. The importance of this factor was investigated by comparing the performance of the control group on the first and last lists of words recalled (i.e., the Equating List vs. the First Hierarchy List). No significant difference was found between the mean numbers of words recalled on these two lists (see Table III). Since practice caused no improvement in the score of the control group, it seems plausible to assume that the superiority of the experimental group over the control group in recalling *test words* of the First Hierarchy List is not due to practice.

TABLE III

COMPARISON OF THE CONTROL GROUP ON THE EQUATING LIST AND THE FIRST HIERARCHY LIST

Lists	No. of Subjects	Mean No. of Words Recalled	S.D.	Student's *t*	*P*-value
Equating	27	5.78	1.12	1.00	>0.30
1st Hierarchy	27	5.48	1.48		

An analysis was made to see whether or not certain words were more often recalled from the First Hierarchy List by either the boys or the girls. It was found that there was no definite preference shown for any one word by either the boys or the girls.

The introspections of the *S*s in response to the questions asked at the end of the experimental session can be briefly summarized. In answer to the first question ("Do you have anything you would like to say about the words?") no significant remarks were volunteered by any *S*. When presented with the second set of questions ("Did anything in particular help you in recalling the words? Did you use any special aids?") a few *S*s said words beginning with the same letter were easy to remember; but the *S*s were, in general, vague about the use of aids and usually replied that they "just remembered, that's all." In response to the last set of questions ("Did you notice any relationship between the words on any of the lists? If so, what was the relationship?") they had much more to say. They all commented upon the fact that the reinforcement lists contained the same words arranged in four different orders. Most of the experimental *S*s were able to indicate that they had been aware of the relationship between the Specific Object List and the test words of the First Hierarchy List. Many of them were not able to verbalize this relationship, but they did give illustrations using words from the two lists; for example, "We had dog and animal, five and number." Probably the best verbalization obtained was offered by a little girl who stated that "an animal is what a dog is."

The generalizing from a lower to a higher level of a hierarchy found in this experiment is another example of semantic or mediated generalization (**1**, **4**). If the associations between species and genus had not been established prior to the experiment, the generalization would not have occurred. The process whereby these associations are set up can be adequately described, for the present at any rate, by general principles of conditioning. For an elaboration of this point the reader is referred to Cofer and Foley (**1**).

In other experiments Long and Welch (**5**, **6**) have seen *S*s of this age range make very effective use of the species-genus relationship in complicated reasoning situations. The results from the previous

and present experiments clearly indicate the ability of such *Ss* to understand and use these relationships.

Summary

This experiment was designed to study the effect of a particular type of verbal reinforcement upon the mediation of generalization. The species-genus relationship was chosen for study. The test words of one list were the generic names of the words on which the *Ss* were trained. Fifty-four children served as *Ss*. After they were tested on an Equating List of words, the children were divided into two groups, equated on the basis of the number of words recalled. The reinforcement list of words (Specific Object List) was shown four times to the members of one group (experimental group), whereas a list of unrelated words (Control List) was shown four times to the members of the other group (control group). The *Ss* in both groups were then presented with the First Hierarchy List (10 generic terms for words on which the experimental group was trained and five unrelated words). Upon the completion of the recall the *Ss* were asked questions designed to elicit verbal comments about any relationship between words of the various lists.

On the basis of the results of this experiment it can be concluded that training on words of a lower hierarchy level (i.e., at the species level) will increase the recall score for words of a higher hierarchy level (i.e., at the genus level). Practice was ruled out as a possible explanation of this finding. Verbal comments in response to specific questions indicated that the species-genus relationship was perceived, although vaguely, by most of the *Ss* in the experimental group.

(Manuscript received June 1, 1944)

References

1. Cofer, C. N., & Foley, J. P., Jr. Mediated generalization and the interpretation of verbal behavior: I. Prolegomena. *Psychol. Rev.*, 1942, **49**, 513–540.
2. Foley, J. P., Jr., & Cofer, C. N. Mediated generalization and the interpretation of verbal behavior: II. Experimental study of certain homophone and synonym gradients. *J. exp. Psychol.*, 1943, **32**, 168–175.
3. Gerbrands, R. Scientific instruments. (Mimeographed)
4. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* N. Y.: Appleton-Century, 1940. Pp. 429.
5. Long, L., & Welch, L. Influence of levels of abstractness on reasoning ability. *J. Psychol.*, 1942, **13**, 41–59.
6. Long, L., & Welch, L. Factors affecting efficiency of inductive reasoning. *J. exp. Educ.*, 1942, **10**, 252–264.
7. Razran, G. H. S. A quantitative study of meaning by a conditioned salivary technique (semantic conditioning). *Science*, 1939, **90**, 89–90.
8. Riess, B. F. Semantic conditioning involving the galvanic skin reflex. *J. exp. Psychol.*, 1940, **26**, 238–240.
9. Thorndike, E. L. *Teacher's word book of 20,000 words.* N. Y.: Teach. Coll. Press, 1932. Pp. 182.

DISCUSSION

AN ALTERNATIVE EXPLANATION OF THE EFFECT OF TEMPERATURE UPON RETENTION IN THE GOLDFISH

BY F. NOWELL JONES

University of Alabama

In this Journal, French (1) has published the results of an investigation of the effect of periods spent in water of different temperatures upon the retention of a maze habit by goldfish. In a carefully conducted experiment, he found that a sojourn in water of 4° C. was followed by a smaller number of relearning errors than was a sojourn in 28° water, and the scores of fish kept in 16° water during the interval (24 hours) between learning and relearning fell between these other two averages. French concludes (1, p. 87):

> Two control experiments were carried out. One of these showed that the results were due to the action of temperature upon retention and not to its action upon the rate of learning. The other showed that the differences were not due to retroactive inhibition produced by general activity during the period of controlled temperature.
>
> It is concluded that there exists a forgetting process which is independent of inhibition and whose rate is a function of temperature.

In reaching these conclusions the factor of habituation or adaptation has been overlooked, and if it is considered, a different hypothesis to account for these results is possible. The evidence for this factor and its relationship to French's results will be discussed below.

Wells (4), Sumner and Wells (3), and Sumner and Doudoroff (2) have published the papers most germane to our discussion. Wells measured the consumption of oxygen by fish (*Gillichthys mirabilis*) in water of 20° and 22° C. after a stay in water of either 11° or 33°. He found that a fish kept at a low temperature consumed more oxygen, i.e., had a higher metabolic rate, when placed in 'neutral' water than did those previously kept at a higher temperature. Wells summarized his results as follows (4, p. 367):

> The results of two experiments are presented showing the effect on the oxygen metabolism produced by acclimatizing fishes to high and low temperatures. It is concluded from an analysis of these data that fishes acclimatized to a high temperature have a much lower rate of metabolism at a common intermediate temperature than do fishes acclimatized to a low temperature, and that the magnitude of this difference is a function of the difference between the acclimatization temperatures.

These effects last for several days. This conclusion was amply supported by the other two papers cited, although somewhat 'oblique' criteria were employed. Sumner and Wells were interested in the effects of drugs and anoxia, as were Sumner and Doudoroff.

Applied to French's experiment, these results make it reasonable to assume that his 4° fish were *metabolizing more rapidly* than the 16° and 28° groups *when returned to 22° water*. To a lesser extent, the 16° group was also metabolizing more rapidly than the 28° group. There is no doubt that the 24-hour (or slightly shorter) period used by French between learning and relearning was sufficient to produce this differential effect (*cf.* Wells and Doudoroff, 2). The better performance (not necessarily better retention) by the 4° group might be explained, therefore, as an accompaniment of a higher metabolic rate, with an associated greater efficiency of the nervous system, and an identical argument may be advanced for the 16° group. An interesting comparison can be made between the error scores for the first relearning day of French's experiment and the metabolic rates found by Wells. French's 4° group made 57 percent as many errors as his 28° group, while the metabolic rate after one hour in 22° water of Wells' 11° fish was 67 percent of that of his 33° fish. Since the difference between 4° and 28° is 24°, and that between 11° and 33° is 22°, the situations are roughly comparable, and the results coincide remarkably well. The suggestion that it is the metabolic rate of the fish at the time of relearning which is determining their error scores is certainly not proved by this coincidence, but nevertheless it is rendered considerably more plausible.

It will be recalled from the summary quoted from French, that two control experiments were conducted. Since these cannot be so simply explained as the primary results which have just been considered, it will be well to discuss them separately. First of all, the activity of fish in water of 4°, 16°, and 28° was measured in order to rule out the possibility that excess movement during the interval between learning and relearning had produced more retroactive inhibition in the fish kept at higher temperatures. Somewhat surprisingly, it was found that fish in 4° water made 1608 'activity units,' those at 16° made 12,764 units, but those at 28° made only 4996 units. It is obvious that if the activity of the fish depended in some simple way upon temperature, the 28° fish should have been most active of all. The experiment of Sumner and Doudoroff showed that respiratory movements at 30° were no greater than at 20°, although the number of such movements at 10° was much smaller. Here, then, is another instance where an overt activity did not follow perfectly the metabolic rate of the fish, since it is definitely proved that oxygen consumption rises with the temperature (Wells, 5; Sumner and Doudoroff, 2). Overt activity must be complexly determined, and perhaps temperatures around 16° are actually stimulating to the receptors in the fish's integument, or they may evoke the release of hormones, or some obscure internal patterns. A more likely explanation would be a simple physical one. The 28° water certainly contained less dissolved oxygen than the 16° water, since, from French's report, no artificial aeration was maintained. The fish may have lacked sufficient oxygen for extended activity. In any event, French is certainly correct in discounting the effect of the interpolated activity, *qua* activity. What effect these various temperatures had upon activity after the fish had returned to 22° is another matter. Data are needed concerning this point. If overt activity behaves as does respiration, a decrease would be expected

for the 28° fish, and an increase for the other two groups (Sumner and Doudoroff, **2**). Such a prediction would be extremely hazardous, however.

The second control experiment involved exposing fish to water of the three experimental temperatures—4°, 16°, and 28°—before original learning. It was found that, under these circumstances, the 4° group learned with an average of 39.9 errors, the 16° group with an average of 25.3 errors, and the 28° group with an average of 29.2 errors. These are very interesting results. If these scores are compared with the error scores made by the experimental groups on the day of original learning, only the 4° group, with its average of 39.3 errors, corresponds to the approximately 42 errors made by the experimental fish. There is no reason apparent why a sojourn at 16° should increase the learning efficiency of a fish, while at the same time treatment by chilling to 4° has no effect. This result is rendered even more difficult of interpretation by the fact that a stay in 28° water has much the same evident effect as the stay at 16°. And both treatments increase learning efficiency over that shown by fish which have presumably been kept at 22°.[1] All that can be said is that the 4° fish were undoubtedly metabolizing more rapidly than those of the other two groups. Perhaps a higher metabolic rate is conducive to efficient performance of a learned act, but has little effect upon original learning (39 vs. 42 errors), while a somewhat lowered metabolic rate promotes original learning. But even then there is this inversion of the scores of the 16° and the 28° groups to be explained. The guess might be hazarded that the 28° group made more errors than the 16° group in original learning, even though metabolizing more slowly, because of some deleterious after-effect of exposure to the higher temperature. Raising them to 28° from 22° risked the supersaturation of the water with dissolved gases, and possibly some effect such as 'bubble-disease' therefrom. An alternative suggestion would be that the particular fish used in this control experiment were, for some reason, better learners than the original group, and the 16° and 28° fish were simply showing no effect from their treatment, while the 4° fish were injured by theirs. If the temperature at which they had been kept by the dealer were known, and their precise history in the laboratory were available, this possibility could be evaluated. If the fish had been kept for some time at high temperatures (or higher than in the case of the experimental groups, at least), they would have been very sensitive to the effects of very low temperatures, and thus the 4° group might have suffered actual injury. The likelihood of injury was certainly increased by the sudden change from 4° to 22°. A change of 18° in 15 minutes is probably too rapid for the welfare of the fish.

The discussion above has been presented primarily to suggest the possibility of an alternative explanation for French's results. It emphasizes the importance of the level of metabolic activity of the fish at the time of relearning, rather than postulating a process of decay during the interval. Obviously, more research would be necessary to determine the exact nature of these processes. In any event, the discussion points out that the state of

[1] French suggests that a difference in stock may have been the cause of this difference. More likely it was the time of year, since metabolic activity varies considerably with the season—even in water of the same temperature (*cf.* Wells, **5**).

metabolism in the fish is *not* conditioned merely by the temperature of the moment, and therefore performance in a maze may *not* be explained as caused solely by the effects of the temperature at any moment upon the bodily chemistry of the fish. French's postulation of "a forgetting process which is independent of inhibition and whose rate is a function of temperature" results from an oversimplification of the factors actually involved, and is not warranted by the data so far available.

(Manuscript received June 26, 1944)

Bibliography

1. French, J. W. The effect of temperature on the retention of a maze habit in fish. *J. exp. Psychol.*, 1942, **31**, 79–87.
2. Sumner, F. B., & Doudoroff, P. Some experiments upon temperature acclimatization and respiratory metabolism in fishes. *Biol. Bull.*, 1938, **74**, 403–429.
3. Sumner, F. B., & Wells, N. A. Some relations between respiratory metabolism in fishes and susceptibility to certain anesthetics and lethal agents. *Biol. Bull.*, 1935, **69**, 368–378.
4. Wells, N. A. Change in rate of respiratory metabolism in a teleost fish induced by acclimatization to high and low temperature. *Biol. Bull.*, 1935, **69**, 361–367.
5. Wells, N. A. The influence of temperature upon the respiratory metabolism of the Pacific killifish, Fundulus parvipinnis. *Physiol. Zoöl.*, 1935, **8**, 196–227.